RELATIVE TAIL ENTROPY FOR RANDOM BUNDLE TRANSFORMATIONS

Xianfeng Ma

Department of Mathematics
East China University of Science and Technology, Shanghai 200237, China

Ercai Chen

School of Mathematical Science
Nanjing Normal University, Nanjing 210097, China
and
Center of Nonlinear Science
Nanjing University, Nanjing 210093, China

Abstract. We introduce the relative tail entropy to establish a variational principle for continuous bundle random dynamical systems. We also show that the relative tail entropy is conserved by the principal extension.

1. **Introduction.** The entropy measures the complexity of a dynamical systems both in the topological and measure-theoretic settings. The topological entropy measures the maximal dynamical complexity versus an average complexity reflected by the measure-theoretic entropy. The relationship between these two kinds of entropy is the classical variational principle, which states that the topological entropy is the supremum of the measure-theoretic entropy over all invariant measures [13, 14, 24].

The entropy concepts can be localized by defining topological tail entropy to quantify the amount of disorder or uncertainty in a system at arbitrary small scales [23]. The local complexity of a dynamical system can also be measured by the defect of uniformity in the convergence of the measure-theoretic entropy function. A variational principle related these two aspects is established in the case of homeomorphism from subtle results in the theory of entropy structure by Downarowicz [8, 3]. An elementary proof of this variational principle for continuous transformations is obtained in terms of essential partitions by Burguet [4]. Ledrappier [21] presents a variational principle between the topological tail entropy and the defect of upper semi-continuity of the measure-theoretic entropy on the cartesian square of the dynamical system, and prove that topological tail entropy is an invariant under any principal extension.

Kifer and Weiss [18] introduce the relative tail entropies for continuous bundle RDSs by investigating the open covers and spanning subsets and deduce the equivalence between the two notions. It is shown in [19] that the defects of the upper semi-continuity of the relative measure-theoretic entropies are bounded from above by the relative tail entropy.

In this paper we devote to proposing a relative variational principle for the relative tail entropy introduced by using open random covers, which enable us to treat different fibers with different open covers. We also introduce the factor transformation and consider its basic properties related to the invariant measure and the upper semi-continuity of the relative measure-theoretic entropy for continuous bundle RDSs. For the product RDS generated by a given RDS and any other RDS with the same probability space, we obtain a variational inequality, which shows that the defect of the upper semi-continuity of the relative measure-theoretic entropy of any invariant measure in the product RDS cannot exceed the relative tail entropy of the original RDS. When the two continuous bundle RDSs coincide, we construct a maximal invariant measure to ensure that the relative tail entropy could be reached, and establish the variational principle. For the probability space being trivial, it reduces to the variational principle deduced by Ledrappier [21] in deterministic dynamical systems. As an application of the variational principle we show that the relative tail entropy is conserved by any principal extension.

The paper is organized as follows. In Section 2, we recall some background in the ergodic theory, introduce the relative tail entropy with respect to open random covers and state our main results. In Section 3, we give some basic properties of the relative entropy and the relative tail entropy. In Section 4, we devote to the proof of the variational principle and show that the relative tail entropy is an invariant under principal extensions.

2. **Preliminaries and main results.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete countably generated probability space together with a $\mathbb{P}$-preserving transformation ϑ and $(X, \mathcal{B})$ be a compact metric space with the Borel σ-algebra $\mathcal{B}$. Let $\mathcal{E}$ be a measurable subset of $\Omega \times X$ with respect to the product σ-algebra $\mathcal{F} \times \mathcal{B}$ and the fibers $\mathcal{E}_\omega = \{x \in X : (\omega, x) \in \mathcal{E}\}$ be compact. A continuous bundle random dynamical system (RDS) T over $(\Omega, \mathcal{F}, \mathbb{P}, \vartheta)$ is generated by the mappings $T_\omega : \mathcal{E}_\omega \to \mathcal{E}_{\vartheta\omega}$ so that the map $(\omega, x) \to T_\omega x$ is measurable and the map $x \to T_\omega x$ is continuous for $\mathbb{P}$-almost all (a.a.) ω. The family $\{T_\omega : \omega \in \Omega\}$ is called a random transformation and each T_ω maps the fiber $\mathcal{E}_\omega$ to $\mathcal{E}_{\vartheta\omega}$. The map $\Theta : \mathcal{E} \to \mathcal{E}$ defined by $\Theta(\omega, x) = (\vartheta\omega, T_\omega x)$ is called the skew product transformation. Observe that $\Theta^n(\omega, x) = (\vartheta^n\omega, T_\omega^n x)$, where $T_\omega^n = T_{\vartheta^{n-1}\omega} \circ \cdots T_{\vartheta\omega} \circ T_\omega$ for $n \geq 0$ and $T_\omega^0 = id$.

Let $\mathcal{P}_\mathbb{P}(\Omega \times X)$ be the space of probability measures on $\Omega \times X$ having the marginal $\mathbb{P}$ on Ω and set $\mathcal{P}_\mathbb{P}(\mathcal{E}) = \{\mu \in \mathcal{P}_\mathbb{P}(\Omega \times X) : \mu(\mathcal{E}) = 1\}$. Denote by $\mathcal{I}_\mathbb{P}(\mathcal{E})$ the space of all $\Theta-$invariant measures in $\mathcal{P}_\mathbb{P}(\mathcal{E})$.

Let $\mathcal{S}$ be a sub-σ-algebra of $\mathcal{F} \times \mathcal{B}$ restricted on $\mathcal{E}$, $\mathcal{R} = \{R_i\}$ be a finite or countable partition of $\mathcal{E}$ into measurable sets. For $\mu \in \mathcal{P}_\mathbb{P}(\Omega \times X)$ the conditional entropy of $\mathcal{R}$ given σ-algebra $\mathcal{S}$ is defined as

$$H_\mu(\mathcal{R} \mid \mathcal{S}) = -\int \sum_i E(1_{R_i} \mid \mathcal{S}) \log E(1_{R_i} \mid \mathcal{S}) d\mu,$$

where $E(1_{R_i} \mid \mathcal{S})$ is the conditional expectation of 1_{R_i} with respect to $\mathcal{S}$.

Let $\mu \in \mathcal{I}_{\mathbb{P}}(\mathcal{E})$ and $\mathcal{S}$ is a sub$-\sigma-$algebra of $\mathcal{F} \times \mathcal{B}$ restricted on $\mathcal{E}$ satisfying $\Theta^{-1}\mathcal{S} \subset \mathcal{S}$. For a given measurable partition $\mathcal{R}$ of $\mathcal{E}$, the conditional entropy $H_{\mu}(\mathcal{R}^{(n)} \mid \mathcal{S})$ is a non-negative sub-additive sequence, where $\mathcal{R}^{(n)} = \bigvee_{i=0}^{n-1}(\Theta^{i})^{-1}\mathcal{R}$. The *relative entropy* $h_{\mu}(\mathcal{R} \mid \mathcal{S})$ *of* Θ *with respect to a partition* $\mathcal{R}$ is defined as

$$h_{\mu}(\mathcal{R} \mid \mathcal{S}) = \lim_{n \to \infty} \frac{1}{n} H_{\mu}(\mathcal{R}^{(n)} \mid \mathcal{S}) = \inf_{n} \frac{1}{n} H_{\mu}(\mathcal{R}^{(n)} \mid \mathcal{S}).$$

The *relative entropy of* Θ is defined by the formula

$$h_{\mu}(\Theta \mid \mathcal{S}) = \sup_{\mathcal{R}} h_{\mu}(\mathcal{R} \mid \mathcal{S}),$$

where the supremum is taken over all finite or countable measurable partitions $\mathcal{R}$ of $\mathcal{E}$ with finite conditional entropy $H_{\mu}(\mathcal{R} \mid \mathcal{S}) < \infty$. The *defect of upper semicontinuity of the relative entropy* $h_{\mu}(\Theta \mid \mathcal{S})$ is defined on $\mathcal{I}_{\mathbb{P}}(\mathcal{E})$ as

$$h_{m}^{*}(\Theta \mid \mathcal{S}) = \begin{cases} \limsup_{\mu \to m} h_{\mu}(\Theta \mid \mathcal{S}) - h_{m}(\Theta \mid \mathcal{S}), & \text{if } h_{m}(\Theta \mid \mathcal{S}) < \infty, \\ \infty, & \text{otherwise.} \end{cases}$$

Any $\mu \in \mathcal{P}_{\mathbb{P}}(\mathcal{E})$ on $\mathcal{E}$ disintegrates $d\mu(\omega, x) = d\mu_{\omega}(x)d\mathbb{P}(\omega)$ (see [10, Section 10.2]), where μ_{ω} are regular conditional probabilities with respect to the $\sigma-$algebra $\mathcal{F}_{\mathcal{E}}$ formed by all sets $(F \times X) \cap \mathcal{E}$ with $F \in \mathcal{F}$. This means that μ_{ω} is a probability measure on $\mathcal{E}_{\omega}$ for $\mathbb{P}$-a.a. ω and for any measurable set $R \in \mathcal{E}$, $\mathbb{P}$-a.s. $\mu_{\omega}(R(\omega)) = E(R \mid \mathcal{F}_{\mathcal{E}})$, where $R(\omega) = \{x : (\omega, x) \in R\}$ and so $\mu(R) = \int \mu_{\omega}(R(\omega))d\mathbb{P}(\omega)$. The conditional entropy of $\mathcal{R}$ given $\sigma-$algebra $\mathcal{F}_{\mathcal{E}}$ can be written as

$$H_{\mu}(\mathcal{R} \mid \mathcal{F}_{\mathcal{E}}) = -\int \sum_{i} E(R_i \mid \mathcal{F}_{\mathcal{E}}) \log E(R_i \mid \mathcal{F}_{\mathcal{E}})d\mathbb{P} = \int H_{\mu_{\omega}}(\mathcal{R}(\omega))d\mathbb{P},$$

where $\mathcal{R}(\omega) = \{R_i(\omega)\}$, $R_i(\omega) = \{x \in \mathcal{E}_{\omega} : (\omega, x) \in R_i\}$ is a partition of $\mathcal{E}_{\omega}$.

Let $(Y, \mathcal{C})$ be a compact metric space with the Borel σ-algebra $\mathcal{C}$ and $\mathcal{G}$ be a measurable, with respect to the product σ-algebra $\mathcal{F} \times \mathcal{C}$, subset of $\Omega \times Y$ with the fibers $\mathcal{G}_{\omega}$ being compact. The continuous bundle RDS S over $(\Omega, \mathcal{F}, \mathbb{P}, \vartheta)$ is generated by the mappings $S_{\omega} : \mathcal{G}_{\omega} \to \mathcal{G}_{\vartheta\omega}$ so that the map $(\omega, y) \to S_{\omega}y$ is measurable and the map $y \to S_{\omega}y$ is continuous for $\mathbb{P}$-almost all (a.a.) ω. The skew product transformation $\Lambda : \mathcal{G} \to \mathcal{G}$ is defined as $\Lambda(\omega, y) = (\vartheta\omega, S_{\omega}y)$.

Definition 2.1. Let T, S are two continuous bundle RDSs over $(\Omega, \mathcal{F}, \mathbb{P}, \vartheta)$ on $\mathcal{E}$ and $\mathcal{G}$, respectively. T is said to be a *factor* of S, or that S is an *extension* of T, if there exists a family of continuous surjective maps $\pi_{\omega} : \mathcal{G}_{\omega} \to \mathcal{E}_{\omega}$ such that the map $(\omega, y) \to \pi_{\omega}y$ is measurable and $\pi_{\vartheta\omega}S_{\omega} = T_{\omega}\pi_{\omega}$. The map $\pi : \mathcal{G} \to \mathcal{E}$ defined by $\pi(\omega, y) = (\omega, \pi_{\omega}y)$ is called the *factor or extension transformation* from $\mathcal{G}$ to $\mathcal{E}$. The skew product system $(\mathcal{E}, \Theta)$ is called a *factor* of $(\mathcal{G}, \Lambda)$ or that $(\mathcal{G}, \Lambda)$ is an *extension* of $(\mathcal{E}, \Theta)$.

Denote by $\mathcal{A}$ the restriction of $\mathcal{F} \times \mathcal{B}$ on $\mathcal{E}$ and set $\mathcal{A}_{\mathcal{G}} = \{\pi^{-1}A : A \in \mathcal{A}\}$.

Definition 2.2. A continuous bundle RDS T on $\mathcal{E}$ is called a *principal factor* of S on $\mathcal{G}$, or that S is a *principal extension* of T, if for any $\Lambda-$invariant probability measure m in $\mathcal{I}_{\mathbb{P}}(\mathcal{G})$, the relative entropy of Λ with respect to $\mathcal{A}_{\mathcal{G}}$ vanishes, *i.e.*, $h_{\mu}(\Lambda \mid \mathcal{A}_{\mathcal{G}}) = 0$.

Let T and S are two continuous bundle RDSs over $(\Omega, \mathcal{F}, \mathbb{P}, \vartheta)$ on $\mathcal{E}$ and $\mathcal{G}$, respectively. Let $\mathcal{H} = \{(\omega, y, x) : y \in \mathcal{G}_{\omega}, x \in \mathcal{E}_{\omega}\}$ and $\mathcal{H}_{\omega} = \{(y, x) : (\omega, y, x) \in \mathcal{H}\}$. It is not hard to see that $\mathcal{H}$ is a measurable subset of $\Omega \times Y \times X$ with respect to

the product $\sigma-$algebra $\mathcal{F} \times \mathcal{C} \times \mathcal{B}$ (as a graph of a measurable multifunction; see [5, Proposition III.13]). The continuous bundle RDS $S \times T$ over $(\Omega, \mathcal{F}, \mathbb{P}, \vartheta)$ is generated by the family of mappings $(S \times T)_\omega : \mathcal{H}_\omega \to \mathcal{H}_{\vartheta\omega}$ with $(y, x) \to (S_\omega y, T_\omega x)$. The map $(\omega, y, x) \to (S_\omega y, T_\omega x)$ is measurable and the map $(y, x) \to (S_\omega y, T_\omega x)$ is continuous in (y, x) for $\mathbb{P}$-a.a. ω. The skew product transformation Γ generated by Θ and Λ from $\mathcal{H}$ to itself is defined as $\Gamma(\omega, y, x) = (\vartheta\omega, S_\omega y, T_\omega x)$.

Let $\pi_\mathcal{E} : \mathcal{H} \to \mathcal{E}$ be the natural projection with $\pi_\mathcal{E}(\omega, y, x) = (\omega, x)$, and $\pi_\mathcal{G} : \mathcal{H} \to \mathcal{G}$ with $\pi_\mathcal{G}(\omega, y, x) = (\omega, y)$. Then $\pi_\mathcal{E}$ and $\pi_\mathcal{G}$ are two factor transformations from $\mathcal{H}$ to $\mathcal{E}$ and $\mathcal{G}$, respectively. Denote by $\mathcal{D}$ the restriction of $\mathcal{F} \times \mathcal{C}$ on $\mathcal{G}$ and set $\mathcal{D}_\mathcal{H} = \pi_\mathcal{G}^{-1}(\mathcal{D}) = \{(D \times X) \cap \mathcal{H} : D \in \mathcal{D}\}$, $\mathcal{A}_\mathcal{H} = \pi_\mathcal{E}^{-1}(\mathcal{A}) = \{(A \times Y) \cap \mathcal{H} : A \in \mathcal{A}\}$, and $\mathcal{F}_\mathcal{H} = \{(F \times Y \times X) \cap \mathcal{H} : F \in \mathcal{F}\}$.

The *relative entropy of* Γ *given the* $\sigma-$*algebra* $\mathcal{D}_H$ is defined by

$$h_\mu(\Gamma \mid \mathcal{D}_H) = \sup_\mathcal{R} h_\mu(\mathcal{R} \mid \mathcal{D}_H),$$

where

$$h_\mu(\mathcal{R} \mid \mathcal{D}_H) = \lim_{n \to \infty} \frac{1}{n} H_\mu \left(\bigvee_{i=0}^{n-1} (\Gamma^i)^{-1} \mathcal{R} \mid \mathcal{D}_H \right)$$

is the *relative entropy of* Γ *with respect to a measurable partition* $\mathcal{R}$, and the supremum is taken over all finite or countable measurable partitions $\mathcal{R}$ of $\mathcal{H}$ with finite conditional entropy $H_\mu(\mathcal{R} \mid \mathcal{D}_H) < \infty$.

Let $\mathcal{E}^{(2)} = \{(\omega, x, y) : x, y \in \mathcal{E}_\omega\}$, which is also a measurable subset of $\Omega \times X^2$ with respect to the product $\sigma-$algebra $\mathcal{F} \times \mathcal{B}^2$. Let $\Theta^{(2)} : \mathcal{E}^{(2)} \to \mathcal{E}^{(2)}$ be a skew-product transformation with $\Theta^{(2)}(\omega, x, y) = (\vartheta\omega, T_\omega x, T_\omega y)$. The map $(\omega, x, y) \to (T_\omega x, T_\omega y)$ is measurable and the map $(x, y) \to (T_\omega x, T_\omega y)$ is continuous in (x, y) for $\mathbb{P}$-a.a. ω. Let $\mathcal{E}_1, \mathcal{E}_2$ be two copies of $\mathcal{E}$, *i.e.*, $\mathcal{E}_1 = \mathcal{E}_2 = \mathcal{E}$, and $\pi_{\mathcal{E}_i}$ be the natural projection from $\mathcal{E}^{(2)}$ to $\mathcal{E}_i$ with $\pi_{\mathcal{E}_i}(\omega, x_1, x_2) = (\omega, x_i)$, i=1, 2. Denote by $\mathcal{A}_{\mathcal{E}^{(2)}} = \{(A \times X) \cap \mathcal{E}^{(2)} : A \in \mathcal{F} \times \mathcal{B}\}$.

The *relative entropy of* $\Theta^{(2)}$ *given the* $\sigma-$*algebra* $\mathcal{A}_{\mathcal{E}^{(2)}}$ is defined by

$$h_\mu(\Theta^{(2)} \mid \mathcal{A}_{\mathcal{E}^{(2)}}) = \sup_\mathcal{R} h_\mu(\mathcal{R} \mid \mathcal{A}_{\mathcal{E}^{(2)}}),$$

where

$$h_\mu(\mathcal{R} \mid \mathcal{A}_{\mathcal{E}^{(2)}}) = \lim_{n \to \infty} \frac{1}{n} H_\mu \left(\bigvee_{i=0}^{n-1} ((\Theta^{(2)})^i)^{-1} \mathcal{R} \mid \mathcal{A}_{\mathcal{E}^{(2)}} \right)$$

is the *relative entropy of* $\Theta^{(2)}$ *with respect to a measurable partition* $\mathcal{R}$, and the supremum is taken over all finite or countable measurable partitions $\mathcal{R}$ of $\mathcal{E}^{(2)}$ with finite conditional entropy $H_\mu(\mathcal{R} \mid \mathcal{A}_{\mathcal{E}^{(2)}}) < \infty$.

A (closed) random set Q is a measurable set valued map $\mathcal{Q} : \Omega \to 2^X$, or the graph of Q denoted by the same letter, taking values in the (closed) subsets of compact metric space X. An open random set U is a set valued map $U : \Omega \to 2^X$ whose complement U^c is a closed random set. A measurable set Q is an open (closed) random set if the fiber Q_ω is an open (closed) subset of $\mathcal{E}_\omega$ in its induced topology from X for $\mathbb{P}-$almost all ω (see [6, Lemma 2.7]). A random cover $\mathcal{Q}$ of $\mathcal{E}$ is a finite or countable family of random sets $\{Q\}$ such that $\mathcal{E}_\omega = \bigcup_{Q \in \mathcal{Q}} Q(\omega)$ for all $\omega \in \Omega$, and it will be called an open random cover if all $Q \in \mathcal{Q}$ are open random sets. Set $\mathcal{Q}(\omega) = \{Q(\omega)\}$, $\mathcal{Q}^{(n)} = \bigvee_{i=0}^{n-1} (\Theta^i)^{-1} \mathcal{Q}$ and $\mathcal{Q}^{(n)}(\omega) = \bigvee_{i=0}^{n-1} (T_\omega^i)^{-1} \mathcal{Q}(\vartheta^i \omega)$. Denote by $\mathfrak{P}(\mathcal{E})$ the set of random covers and $\mathfrak{U}(\mathcal{E})$ the set of open random covers.

For $\mathcal{R}, \mathcal{Q} \in \mathfrak{P}(\mathcal{E})$, $\mathcal{R}$ is said to be finer than $\mathcal{Q}$, which we will write $\mathcal{R} \succ \mathcal{Q}$ if each element of $\mathcal{R}$ is contained in some element of $\mathcal{Q}$.

For any non-empty set $S \subset \mathcal{E}$ and a random cover $\mathcal{R} \in \mathfrak{P}(\mathcal{E})$, let $N(S, \mathcal{R}) = \min\{\text{card}(\mathcal{U}) : \mathcal{U} \subset \mathcal{R}, S \subset \cup_{u \in \mathcal{U}} U\}$ and $N(\emptyset, \mathcal{R}) = 1$. Denote by $N(S, \mathcal{R})(\omega) = \min\{\text{card}(\mathcal{U}(\omega)) : \mathcal{U}(\omega) \subset \mathcal{R}(\omega), S(\omega) \subset \cup_{u \in \mathcal{U}} U(\omega)\}$. Clearly, $N(S, \mathcal{R})(\omega) \leq N(S, \mathcal{R})$ for each ω. For $\mathcal{R}, \mathcal{Q} \in \mathfrak{P}(\mathcal{E})$, let $N(\mathcal{R} \mid \mathcal{Q}) = \max_{Q \in \mathcal{Q}} N(Q, \mathcal{R})$ and $N(\mathcal{R} \mid \mathcal{Q})(\omega) = \max_{Q \in \mathcal{Q}} N(Q, \mathcal{R})(\omega)$.

Lemma 2.1. *Let $\mathcal{R} \in \mathfrak{U}(\mathcal{E})$ and $\mathcal{Q} \in \mathfrak{P}(\mathcal{E})$. The function $\omega \to N(\mathcal{R} \mid \mathcal{Q})(\omega)$ is measurable.*

Proof. Let $Q \in \mathcal{Q}$ and $\mathcal{R} = \{R_1, \ldots, R_l\}$. For each ω, there exists a subset $\{j_1, \ldots j_k\}$ of $\{1, \ldots, l\}$ such that $Q(\omega) \subset \cup_{i=1}^{k} R_{j_i}(\omega)$. Let

$$\Omega_{j_1, \ldots, j_k} = \{\omega \in \Omega : Q(\omega) \subset \bigcup_{i=1}^{k} R_{j_i}(\omega)\}.$$

Since Q is a random set and

$$\Omega_{j_1, \ldots, j_k} = \Omega \setminus \{\omega : \left(\mathcal{E} \setminus \bigcup_{i=1}^{k} R_{j_i}(\omega)\right) \cap Q(\omega) \neq \emptyset\},$$

$\Omega_{j_1, \ldots, j_k}$ is a measurable subset of Ω (see for instance [5, Theorem II.30]). One obtain a finite partition of Ω into measurable sets Ω^J, where J is a finite family of subsets of $\{1, \ldots, l\}$ such that $\Omega^J = \bigcap_{(j_1, \ldots, j_k) \in J} \Omega_{j_1, \ldots, j_k}$. Thus for each ω,

$$N(Q, \mathcal{R})(\omega) = \min_{(j_1, \ldots, j_k) \in J, 1 \leq k \leq l} \text{card}\{j_1, \ldots, j_k\},$$

and $N(Q, \mathcal{R})(\omega)$ is measurable in ω.

Notice that for each $t \in \mathbb{R}$,

$$\{\omega : N(\mathcal{R} \mid \mathcal{Q})(\omega) > t\} = \bigcup_{Q \in \mathcal{Q}} \{\omega : N(Q, \mathcal{R})(\omega) > t\}.$$

The result holds from the measurability of $N(Q, \mathcal{R})(\omega)$ in ω. $\square$

For any $\mathcal{R}, \mathcal{Q}, \mathcal{U}, \mathcal{V} \in \mathfrak{P}(\mathcal{E})$, the following inequalities always hold.

$$N(\mathcal{R} \mid \mathcal{Q})(\omega) \leq N(\mathcal{U} \mid \mathcal{V})(\omega), \qquad \text{if } \mathcal{U} \succ \mathcal{R}, \mathcal{Q} \succ \mathcal{V}, \tag{1}$$

$$N(\Theta^{-1}\mathcal{R} \mid \Theta^{-1}\mathcal{Q})(\vartheta\omega) \leq N(\mathcal{R} \mid \mathcal{Q})(\omega), \tag{2}$$

$$N(\mathcal{R} \vee \mathcal{Q} \mid \mathcal{U})(\omega) \leq N(\mathcal{R} \mid \mathcal{U})(\omega) \cdot N(\mathcal{Q} \mid \mathcal{R} \vee \mathcal{U})(\omega), \tag{3}$$

$$N(\mathcal{R} \vee \mathcal{Q} \mid \mathcal{U} \vee \mathcal{V})(\omega) \leq N(\mathcal{R} \mid \mathcal{U})(\omega) \cdot N(\mathcal{Q} \mid \mathcal{V})(\omega). \tag{4}$$

Let $\mathcal{R} \in \mathfrak{U}(\mathcal{E})$ and $\mathcal{Q} \in \mathfrak{P}(\mathcal{E})$. By the inequality (2) and (3) it is easy to see that the sequence $\log N(\mathcal{R}^{(n)} \mid \mathcal{Q}^{(n)})(\omega)$ is subadditive for each ω. By the subadditive ergodic theorem (see [26, 16]) the following limit

$$h_\Theta(\mathcal{R} \mid \mathcal{Q})(\omega) = \lim_{n \to \infty} \frac{1}{n} \log N(\mathcal{R}^{(n)} \mid \mathcal{Q}^{(n)})(\omega)$$

$\mathbb{P}$−almost surely (a.s.) exists and

$$h_\Theta(\mathcal{R} \mid \mathcal{Q}) = \lim_{n \to \infty} \frac{1}{n} \int \log N(\mathcal{R}^{(n)} \mid \mathcal{Q}^{(n)})(\omega) d\mathbb{P} = \int h_\Theta(\mathcal{R} \mid \mathcal{Q})(\omega) d\mathbb{P}.$$

$h_\Theta(\mathcal{R} \mid \mathcal{Q})$ will be called *relative tail entropy of Θ on an open random cover $\mathcal{R}$ with respect to a random cover $\mathcal{Q}$* . If $\mathcal{Q}$ is a trivial random cover, then $h_\Theta(\mathcal{R} \mid \mathcal{Q})$ is

called the *relative topological entropy $h_\Theta^{(r)}(\mathcal{R})$ of Θ with respect to an open random cover $\mathcal{R}$*, by (1),

$$h_\Theta^{(r)}(\mathcal{R}) \geq h_\Theta(\mathcal{R} \mid \mathcal{Q}), \tag{5}$$

for all $\mathcal{Q} \in \mathfrak{P}(\mathcal{E})$.

From (1), one can see that

$$h_\Theta(\mathcal{R} \mid \mathcal{Q}) \leq h_\Theta(\mathcal{U} \mid \mathcal{V}) \quad \text{if } \mathcal{U} \succ \mathcal{R}, \mathcal{Q} \succ \mathcal{V}. \tag{6}$$

Then there exists a limit (finite or infinite) over the directed set $\mathfrak{U}(\mathcal{E})$,

$$h(\Theta \mid \mathcal{Q}) = \lim_{\mathcal{R} \in \mathfrak{U}(\mathcal{E})} h_\Theta(\mathcal{R} \mid \mathcal{Q}) = \sup_{\mathcal{R} \in \mathfrak{U}(\mathcal{E})} h_\Theta(\mathcal{R} \mid \mathcal{Q}),$$

which will be called the *relative tail entropy of Θ with respect to a random cover $\mathcal{Q}$*. By the inequality (6),

$$h(\Theta \mid \mathcal{Q}) \leq h(\Theta \mid \mathcal{V}), \quad \text{if } \mathcal{Q} \succ \mathcal{V},$$

then one can take the limit again

$$h^*(\Theta) = \lim_{\mathcal{Q} \in \mathfrak{P}(\mathcal{E})} h(\Theta \mid \mathcal{Q}) = \inf_{\mathcal{Q} \in \mathfrak{P}(\mathcal{E})} h(\Theta \mid \mathcal{Q}),$$

which is called the *relative tail entropy of Θ*. It follows from the inequality (5) that $h^{(r)}(\Theta) \geq h^*(\Theta)$.

Remark 2.1. For each open cover $\xi = \{A_1, \ldots, A_k\}$ of the compact space X, $\{(\Omega \times A_i) \cap \mathcal{E}\}_{i=1}^k$ naturally form an open random cover of $\mathcal{E}$. The relative tail entropy related with this kind of random cover is discussed under the name of "relative conditional entropy" in [18].

One of our main goals is to establish the following variational inequality, which shows that the defect of upper semi-continuity of the relative measure-theoretical entropy function cannot exceed the relative tail entropy.

Theorem 2.1. *Let $S \times T$ be the continuous bundle RDS on $\mathcal{H}$ and $m \in \mathcal{I}_\mathbb{P}(\mathcal{H})$. Then $h_m^*(\Gamma \mid \mathcal{D}_\mathcal{H}) \leq h^*(\Theta)$.*

Remark 2.2. For the trivial space $(Y, \mathcal{C})$ and the random cover mentioned in Remark 2.1, the above result reduces to the theorem presented by Kifer and Liu (See [19, Theorem 1.3.5])

We will obtain the following variational principle when we consider the continuous bundle RDS $T \times T$.

Theorem 2.2. *Let T be a continuous bundle RDS on $\mathcal{E}$. Then*

$$\max\{h_\mu^*(\Theta^{(2)} \mid \mathcal{A}_{\mathcal{E}^{(2)}}) : \mu \in \mathcal{I}_\mathbb{P}(\mathcal{E}^{(2)})\} = h^*(\Theta).$$

Definition 2.3. A continuous bundle RDS T is called *relatively asymptotically h-expansive* if the relative tail entropy $h^*(\Theta) = 0$.

Remark 2.3. Theorem 2.2 indicates that the upper semi-continuity of the function $h_{(\cdot)}(\Theta^{(2)} \mid \mathcal{A}_{\mathcal{E}^{(2)}})$ on $\mathcal{I}_\mathbb{P}(\mathcal{E}^{(2)})$ is equivalent to relatively asymptotically h-expansiveness of T. Moreover, by Theorem 2.1, for a continuous bundle RDS $S \times T$ on $\mathcal{H}$ generated by the continuous bundle RDS T and any other continuous bundle RDS S, relatively asymptotically h-expansiveness of T is also equivalent to the upper semi-continuity of the function $h_{(\cdot)}(\Gamma \mid \mathcal{D}_\mathcal{H})$ on $\mathcal{I}_\mathbb{P}(\mathcal{H})$. In general, the upper semi-continuity of the usual measure-theoretical entropy does not imply the relatively asymptotically h-expansiveness of a random transformation, even in the

deterministic case (see [23, Example 6.4]). An equivalence condition with respect to the upper semi-continuity of the measure-theoretic entropy is given by making use of the local entropy theory (See [15, Lemma 6.4])

As an application of the variational principle, we will derive the following result.

Theorem 2.3. *Let T, S be two continuous bundle RDSs on $\mathcal{E}$ and $\mathcal{G}$, respectively. Suppose that S is a principal extension of T via the factor transformation π, then $h^*(\Lambda) = h^*(\Theta)$.*

Remark 2.4. Theorem 2.3 shows that the relative tail entropy for random transformations could be conserved by the principal extension. If two continuous bundle RDSs have a common principal extension, they are equivalent in the sense of the principal extension.

3. **Relative tail entropy and relative entropy.** we will first give two propositions regarded as the relative tail entropy, which will be needed in the proof of variational inequality later.

Proposition 3.1. *Let T be a continuous bundle RDS on $\mathcal{E}$, and $\mathcal{Q}$ be a random cover of $\mathcal{E}$. Then for each $m \in \mathbb{N}$,*

$$h(\Theta^m \mid \mathcal{Q}^{(m)}) = mh(\Theta \mid \mathcal{Q}),$$

where $\mathcal{Q}^{(m)} = \bigvee_{i=0}^{m-1} (\Theta^i)^{-1}\mathcal{Q}$.

Proof. Let $\mathcal{R}$ be an open random cover of $\mathcal{E}$. Since

$$\bigvee_{j=0}^{n-1} (\Theta^{mj})^{-1}\Big(\bigvee_{i=0}^{m-1} (\Theta^i)^{-1}\mathcal{R}\Big) = \bigvee_{i=0}^{nm-1} (\Theta^i)^{-1}\mathcal{R},$$

Then for each $\omega \in \Omega$,

$$N\Big(\bigvee_{j=0}^{n-1} (\Theta^{mj})^{-1}\mathcal{R}^{(m)}\Big) \mid \bigvee_{j=0}^{n-1} (\Theta^{mj})^{-1}\mathcal{Q}^{(m)})(\omega) = N(\mathcal{R}^{(nm)} \mid \mathcal{Q}^{(nm)})(\omega).$$

By the definition of the relative tail entropy of Θ^m on open random cover $\mathcal{R}^{(m)}$ with respect to $\mathcal{Q}^{(m)}$,

$$h_{\Theta^m}(\mathcal{R}^{(m)} \mid \mathcal{Q}^{(m)}) = \lim_{n \to} \frac{1}{n} \int \log N\Big(\bigvee_{j=0}^{n-1} (\Theta^{mj})^{-1}\mathcal{R}^{(m)}\Big) \mid \bigvee_{j=0}^{n-1} (\Theta^{mj})^{-1}\mathcal{Q}^{(m)})(\omega)d\mathbb{P}$$

$$= \lim_{n \to} \frac{1}{n} \int \log N(\mathcal{R}^{(nm)} \mid \mathcal{Q}^{(nm)})(\omega)d\mathbb{P}$$

$$= \lim_{n \to} m\frac{1}{nm} \int \log N(\mathcal{R}^{(nm)} \mid \mathcal{Q}^{(nm)})(\omega)d\mathbb{P}$$

$$= mh_{\Theta}(\mathcal{R} \mid \mathcal{Q}).$$

Then

$$mh(\Theta \mid \mathcal{Q}) = \sup_{\mathcal{R}} h_{\Theta^m}(\mathcal{R}^{(m)} \mid \mathcal{Q}^{(m)}) \leq h(\Theta^m \mid \mathcal{Q}^{(m)}),$$

where the supremum is taken over all open random covers $\mathcal{R}$ of $\mathcal{E}$.

Since $\mathcal{R} \prec \mathcal{R}^{(m)}$, then by the inequality (1),

$$N\Big(\mathcal{R}^{(m)} \mid \bigvee_{j=0}^{n-1} (\Theta^{mj})^{-1}\mathcal{Q}^{(m)}\Big)(\omega) \leq N\Big(\bigvee_{j=0}^{n-1} (\Theta^{mj})^{-1}\mathcal{R}^{(m)}\Big) \mid \bigvee_{j=0}^{n-1} (\Theta^{mj})^{-1}\mathcal{Q}^{(m)})(\omega),$$

which implies that

$$h_{\Theta^m}(\mathcal{R} \mid \mathcal{Q}^{(m)}) \leq h_{\Theta^m}(\mathcal{R}^{(m)} \mid \mathcal{Q}^{(m)}) = mh_{\Theta}(\mathcal{R} \mid \mathcal{Q}).$$

Thus $h(\Theta^m \mid \mathcal{Q}^{(m)}) \leq mh(\Theta \mid \mathcal{Q})$ and the proposition is proved. $\qquad\square$

We could deduce from Proposition 3.1 the following power rule for the relative tail entropy.

Proposition 3.2. *Let T be a continuous bundle RDS on $\mathcal{E}$. Then for each $m \in \mathbb{N}$,*

$$h^*(\Theta^m) = mh^*(\Theta).$$

Proof. By Proposition 3.1,

$$\inf_{\mathcal{Q}} h(\Theta^m \mid \mathcal{Q}^{(m)}) = \inf_{\mathcal{Q}} mh(\Theta \mid \mathcal{Q}) = mh^*(\Theta),$$

where the infimum is taken over all random covers of $\mathcal{E}$. Then $h^*(\Theta^m) \leq mh^*(\Theta)$.
    Since $\mathcal{Q} \prec \mathcal{Q}^{(m)}$, then

$$h(\Theta^m \mid \mathcal{Q}) \geq h(\Theta^m \mid \mathcal{Q}^{(m)}) \geq mh^*(\Theta).$$

By taking infimum on the inequality over all random covers of $\mathcal{E}$, one get $h^*(\Theta^m) \geq mh^*(\Theta)$ and the equality holds. $\qquad\square$

Let $\mu \in \mathcal{P}_{\mathbb{P}}(\mathcal{E})$. A partition $\mathcal{P}$ is called $\delta-$contains a partition $\mathcal{Q}$ if there exists a partition $\mathcal{R} \preceq \mathcal{P}$ such that $\inf \sum_i \mu(R_i^* \triangle Q_i^*) < \delta$, where the infimum is taken over all ordered partitions $\mathcal{R}^*, \mathcal{Q}^*$ obtained from $\mathcal{R}$ and $\mathcal{Q}$.
    The following lemma essentially comes from the argument of Theorem 4.18 in [25] and Lemma 4.15 in [26].

Lemma 3.1. *Given $\epsilon > 0$ and $k \in \mathbb{N}$. There exists $\delta = \delta(\epsilon, k) > 0$ such that if the measurable partition $\mathcal{P}$ $\delta-$contains $\mathcal{Q}$, where $\mathcal{Q}$ is a finite measurable partition with k elements, then $H_\mu(\mathcal{Q} \mid \mathcal{P}) < \epsilon$.*

Proof. Let $\epsilon > 0$. Choose $0 < \delta < \frac{1}{e}$ such that $-\delta \log \delta + (1-\delta) \log(1-\delta) + \delta \log k < \epsilon$. Suppose that $\mathcal{R} \preceq \mathcal{P}$ is the partition with $\sum_i \mu(R_i \triangle Q_i) < \delta$. One can construct a partition $\mathcal{S}$ by $S_0 = \bigcup_i (R_i \cap Q_i)$ and $S_i = Q_i \setminus S_0$. Since $\mathcal{R} \vee \mathcal{Q} = \mathcal{R} \vee \mathcal{S}$, and

$$H_\mu(\mathcal{R}) + H_\mu(\mathcal{Q} \mid \mathcal{R}) = H_\mu(\mathcal{R} \vee \mathcal{Q}) = H_\mu(\mathcal{R} \vee \mathcal{S}) \leq H_\mu(\mathcal{S}) + H_\mu(\mathcal{R}).$$

Then

$$H_\mu(\mathcal{Q} \mid \mathcal{R}) \leq H_\mu(\mathcal{S}) \leq -\delta \log \delta + (1 - \delta) \log(1 - \delta) + \delta \log k < \epsilon,$$

and $H_\mu(\mathcal{Q} \mid \mathcal{P}) < H_\mu(\mathcal{Q} \mid \mathcal{R}) < \epsilon$. $\qquad\square$

Remark 3.1. We discuss here the conditional entropy instead of the usual measure-theoretic entropy in [25]. The result does not require that the two partitions have the same cardinality, which is a little different from Lemma 4.15 in [26].

Lemma 3.2. *Let $\mu^{(i)} \in \mathcal{P}_{\mathbb{P}}(\mathcal{E}), i \in \mathbb{N}$ and $\delta = \delta(\omega)$ be a positive random variable on Ω. There exists a finite measurable partition $\mathcal{R} = \{R\}$ of $\mathcal{E}$ such that $\operatorname{diam} R(\omega) \leq \delta(\omega)$ $\mathbb{P}-a.s.$ and $\mu^{(i)}(\partial R) = 0$ for each $i \in \mathbb{N}, R \in \mathcal{R}$ in the sense of $\mu^{(i)}(\partial R) = \int \mu_\omega^{(i)}(\partial R(\omega)) d\mathbb{P}(\omega)$, where ∂ denotes the boundary.*

Proof. Since $(\Omega, \mathcal{F}, \mathbb{P})$ is a Lebesgue space, it can be viewed as a Borel subset of the unit interval $[0, 1]$, and $\mu^{(i)}, i \in \mathbb{N}$ are also probability measures on the compact space $[0, 1] \times X$ with the marginal $\mathbb{P}$ on $[0, 1]$.

Fix a point $(t, x) \in [0, 1] \times X$ and $\mu^{(i)} \in \mathcal{P}_\mathbb{P}(\mathcal{E})$. For each nonrandom $\epsilon > 0$, $\partial B((t, x), \epsilon) \subset \overline{B}((t, x), \epsilon) \backslash \text{int} B((t, x), \epsilon)$, where $B((t, x), \epsilon)$ is the open ball of center at (t, x) and radius ϵ with the product metric $d = (d_1^2 + d_2^2)^{\frac{1}{2}}$ on $[0, 1] \times X$ and $\overline{B}$ denotes the closure of B. If $\epsilon_1 \neq \epsilon_2$, $(\overline{B}((t, x), \epsilon_1) \backslash \text{int} B((t, x), \epsilon_1)) \cap (\overline{B}((t, x), \epsilon_2) \backslash \text{int} B((t, x), \epsilon_2)) = \emptyset$. Then there exists only at most countably many of ϵ_j such that $\mu^{(i)}(\overline{B}((t, x), \epsilon_j) \backslash \text{int} B((t, x), \epsilon_j)) > 0, j = 1, 2, \dots$. It follows that for each $(t, x) \in [0, 1] \times X$ one could choose some positive real number $\gamma(t) < \frac{\delta(t)}{2}$ such that for all $i \in \mathbb{N}$, $\mu^{(i)}(\overline{B}((t, x), \gamma(t)) \backslash \text{int} B((t, x), \gamma(t))) = 0$ and then $\mu^{(i)}(\partial B((t, x), \gamma(t))) = 0$.

By the compactness of $[0, 1] \times X$, there exists an open cover of $[0, 1] \times X$ by finite many open ball $B_1, \dots, B_k$ with the diameter of the $t-$section of B_j diam $B_j(t) < \delta(t)$ and $\mu^{(i)}(\partial B_j) = \int \mu_\omega^{(i)} \partial B_j(\omega) d\mathbb{P} = 0$ for each $1 \leq j \leq k$ and $i \in \mathbb{N}$. Then $B_1, B_2 \backslash B_1, \dots, B_k \backslash \cup_{j=1}^{k-1} B_j$ forms a measurable partition of $[0, 1] \times X$. Let $A_1 = B_1$, and $A_n = B_n \backslash \cup_{j=1}^{n-1} B_j$ for each $1 \leq n \leq k$ and denote by $\xi = \{A_1, \dots, A_k\}$. Then ξ is a measurable partition of $[0, 1] \times X$ and for each $1 \leq j \leq k$, diam $A_j(t) < \delta(t)$ for each $t \in [0, 1]$. Since $\partial A_n \subset \cup_{j=1}^{n} \partial B_j$, then $\mu^{(i)}(\partial A_j) = 0$, for each $1 \leq j \leq k$ and $i \in \mathbb{N}$.

Let $R_j = A_j \cap \mathcal{E}$ and $R_j(\omega) = \{x : (\omega, x) \in R_j\}$. Notice that the marginal $\mathbb{P}$ is supported on Ω, then $\mu^{(i)}(\partial R_j) = 0$ and $\mathcal{R} = \{R_1, \dots, R_k\}$ is the measurable partition as desired. $\qed$

Let T and S be two continuous bundle RDSs on $\mathcal{E}$ and $\mathcal{G}$, respectively. Let T be a factor of S via the factor transformation π. The transformation π induce a map, which is again denoted by π, from $\mathcal{P}_\mathbb{P}(\mathcal{G})$ to $\mathcal{P}_\mathbb{P}(\mathcal{E})$ by $\pi\mu = \mu\pi^{-1}$. This induced map π transports every measure on $\mathcal{G}$ to a measure $\pi\mu$ in $\mathcal{P}_\mathbb{P}(\mathcal{E})$. The following proposition is classical in the deterministic dynamical system [7].

Proposition 3.3. *The induced map π is a continuous affine map from $\mathcal{P}_\mathbb{P}(\mathcal{G})$ onto $\mathcal{P}_\mathbb{P}(\mathcal{E})$.*

Proof. If $\mu_n \to \mu \in \mathcal{P}_\mathbb{P}(\mathcal{G})$, then $\int f d\mu_n \to \int f d\mu$ for all $f \in \mathcal{C}(\mathcal{G})$, where $\mathcal{C}(\mathcal{G})$ is the set of random continuous functions on $\mathcal{G}$ (see [6]), and therefore $\int g \circ \pi d\mu_n \to \int g \circ \pi d\mu$ for all $g \in \mathcal{C}(\mathcal{G})$ by the measurability of the factor transformation π. This implies $\pi\mu_n \to \pi\mu$.

It is clear that $\pi(\alpha\mu + (1 - \alpha)\nu) = \alpha\pi\mu + (1 - \alpha)\pi\nu$ for all $\mu, \nu \in \mathcal{P}_\mathbb{P}(\mathcal{G})$ and $0 \leq \alpha \leq 1$.

Since $\mathcal{P}_\mathbb{P}(\mathcal{G})$ is a compact and convex subset of $\mathcal{P}_\mathbb{P}(\Omega \times Y)$ (see [1, Section 1.5]), and the ergodic measures on $\mathcal{G}$ are just the point measures if we take the identity transformation on $\mathcal{G}$, *i.e.*, $Id(\omega, x) = (\omega, x)$. Then by the Krein-Millman theorem (see [11, P440]), the convex combinations of point measures are dense in $\mathcal{P}_\mathbb{P}(\mathcal{G})$. it follows that $\pi : \mathcal{G} \to \mathcal{E}$ is onto that $\pi : \mathcal{P}_\mathbb{P}(\mathcal{G}) \to \mathcal{P}_\mathbb{P}(\mathcal{E})$ is onto. $\qed$

It is not hard to verify that the induced map π send a Λ-invariant measure in $\mathcal{I}_\mathbb{P}(\mathcal{G})$ to a Θ-invariant measure in $\mathcal{I}_\mathbb{P}(\mathcal{E})$. The following result shows that this map is also surjective.

Proposition 3.4. *Let $\mu \in \mathcal{I}_\mathbb{P}(\mathcal{E})$. There exists $m \in \mathcal{I}_\mathbb{P}(\mathcal{G})$ with $\pi m = \mu$*

Proof. By Proposition 3.3, there exists a $\nu \in \mathcal{P}_{\mathbb{P}}(\mathcal{G})$ such that $\pi\nu = \mu$. Since $\Theta\mu = \mu$ and $\pi\Lambda = \Theta\pi$, one has $\pi(\Lambda\nu) = \mu$, and more generally, $\pi(\Lambda^n\nu) = mu$. By the affinity of π, $\pi(\frac{1}{n}\sum_{i=0}^{n-1}\Lambda^i\nu) = \mu$. Denote by $\nu^{(n)} = \frac{1}{n}\sum_{i=0}^{n-1}\Lambda^i\nu$ and let m be one limit point of the sequence $\nu^{(n)}$, It follows from Theorem 1.5.8 in [1] that $m \in \mathcal{I}_{\mathbb{P}}(\mathcal{G})$. Since π is continuous, then $\pi m = \mu$. $\square$

We need the following lemma (see [12, Section 14.3]) which follows from the martingale convergence theorem.

Lemma 3.3. *Let $\mu \in \mathcal{P}_{\mathbb{P}}(\mathcal{G})$, $\mathcal{R} = \{R_1, \ldots, R_k\}$ be a finite measurable partition of $\mathcal{G}$ with $H_\mu(\mathcal{R}) < \infty$ and $\mathcal{A}_1 \prec \cdots \prec \mathcal{A}_n \prec \cdots$ be an increasing sequence of sub-σ-algebra of $\mathcal{A}$ with $\bigvee_{n=1}^\infty = \mathcal{A}$. Then*

$$H_\mu(\mathcal{R} \mid \mathcal{A}_\mathcal{G}) = \lim_{n\to\infty} H_\mu(\mathcal{R} \mid \mathcal{A}_n) = \inf_n H_\mu(\mathcal{R} \mid \mathcal{A}_n).$$

The following result is a relative version of Lemma 6.6.7 in [9]. Similar results for random transformations could be found in [20, 17]

Lemma 3.4. *Let $(\mathcal{E}, \Theta)$ be a factor of $(\mathcal{G}, \Lambda)$ via a factor transformation π, $m \in \mathcal{P}_{\mathbb{P}}(\mathcal{G})$ and $\mathcal{R} = \{R\}$ be a finite measurable partition of $\mathcal{G}$ with $m(\partial R) = 0$, where ∂ denotes the boundary and $m(\partial R) = \int m_\omega(\partial R(\omega))d\mathbb{P}$. Then*

(i) m is a supper semi-continuity point of the function $\mu \to H_\mu(\mathcal{R} \mid \mathcal{A}_\mathcal{G})$ defined on $\mathcal{P}_{\mathbb{P}}(\mathcal{G})$, i.e.,

$$\limsup_{\mu\to m} H_\mu(\mathcal{R} \mid \mathcal{A}_\mathcal{G}) \le H_m(\mathcal{R} \mid \mathcal{A}_\mathcal{G}).$$

(ii) If $m \in \mathcal{I}_{\mathbb{P}}(\mathcal{G})$, the function $\mu \to h_\mu(\mathcal{R} \mid \mathcal{A}_\mathcal{G})$ defined on $\mathcal{I}_{\mathbb{P}}(\mathcal{G})$ is upper semi-continuous at m, i.e.,

$$\limsup_{\mu\to m} h_\mu(\mathcal{R} \mid \mathcal{A}_\mathcal{G}) \le h_m(\mathcal{R} \mid \mathcal{A}_\mathcal{G}).$$

Proof. (i) For $R \in \mathcal{R}$ with $R(\omega) = \{x : (\omega, x) \in R\}$. Let $\overline{R} = \{(\omega, x) : x \in \overline{R(\omega)}\}$ and $\underline{R} = \{(\omega, x) : x \in \text{int}(R(\omega))\}$, where $\overline{R(\omega)}$ and $\text{int}(R(\omega))$ denotes the closure and the interior of $R(\omega)$, respectively. Then $\overline{R}$ is a closed random set of $\mathcal{G}$ and $\underline{R}$ is an open random set. By Portmenteau theorem (see [6]),

$$m(\overline{R}) \ge \limsup_{\mu\to m}\mu(\overline{R}) \ge \limsup_{\mu\to m}\mu(R) \ge \liminf_{\mu\to m}\mu(R) \ge \liminf_{\mu\to m}\mu(\underline{R}) \ge m(\underline{R}).$$

Since $m(\overline{R}) = m(R) = m(\underline{R})$ by $m(\partial R) = 0$, then $\mu \to \mu(R)$ defined on $\mathcal{P}_{\mathbb{P}}(\mathcal{G})$ is continuous at m. Recall that the function $t \to -t\log t$ is continuous on $[0, 1]$. Then $\mu \to H_\mu(\mathcal{R})$ is also continuous at m on $\mathcal{P}_{\mathbb{P}}(\mathcal{G})$. Moreover, if $\mathcal{Q} = \{Q\}$ is a measurable partition of $\mathcal{G}$ with $m(\partial Q) = 0$ for each $Q \in \mathcal{Q}$, then the conditional entropy $\mu \to H_\mu(\mathcal{R} \mid \mathcal{Q})$ of the partition $\mathcal{R}$ over $\mathcal{Q}$ is continuous at m.

Let $\nu = \pi m$. By Lemma 3.2, there exists a refining sequence of finite measurable partitions $\mathcal{Q}_k = \{Q_{k_i}\}$ of $\mathcal{E}$ satisfying $\nu(\partial Q_{k_i}) = 0$, for each $Q_{k_i} \in \mathcal{Q}_k$, $k = 1, 2, \ldots$. Then $\{\pi^{-1}\mathcal{Q}_k\}$ is a refining sequence of measurable partitions of $\mathcal{G}$, all having the boundary of measure zero at m. It follows that for each $k \in \mathbb{N}$, the function $\mu \to H_\mu(\mathcal{R} \mid \pi^{-1}\mathcal{Q}_k)$ is continuous at m. Notice that $\bigvee_{k=1}^\infty \mathcal{Q}_k = \mathcal{A}$ and $H_\mu(\mathcal{R} \mid \pi^{-1}\mathcal{Q}_k)$ decrease in k. Thus the function $\mu \to \inf_k H_\mu(\mathcal{R} \mid \pi^{-1}\mathcal{Q}_k)$ is upper semi-continuous at m and the property (i) follows from Lemma 3.3.

(ii) Let $n \in \mathbb{N}$. Since the function $\mu \to \frac{1}{n}H_\mu(\mathcal{R}^{(n)} \mid \pi^{-1}\mathcal{Q}_k)$ is also continuous at m for each $k = 1, 2, \ldots$, where $\mathcal{R}^{(n)} = \bigvee_{i=0}^{n-1}(\Lambda^i)^{-1}\mathcal{R}$, then the function $\mu \to \inf_k \frac{1}{n}H_\mu(\mathcal{R}^{(n)} \mid \pi^{-1}\mathcal{Q}_k) = \frac{1}{n}H_\mu(\mathcal{R}^{(n)} \mid \mathcal{A}_\mathcal{G})$ is upper semi-continuous at m.

Therefore the function $\mu \to \inf_n \frac{1}{n} H_\mu(\mathcal{R}^{(n)} \mid \mathcal{A}_\mathcal{G}) = h_\mu(\mathcal{R} \mid \mathcal{A}_\mathcal{G})$ is upper semi-continuous at m and the property (ii) holds. $\qquad \square$

We need the following lemma which shows the basic connection between the relative entropy and relative tail entropy.

Lemma 3.5. *Let S be a continuous bundle RDS on $\mathcal{G}$. Suppose that $\mathcal{R} = \{R\}, \mathcal{Q} = \{Q\}$ are two finite measurable partitions of $\mathcal{G}$ and $\mu \in \mathcal{P}_\mathbb{P}(\mathcal{G})$, then*

$$H_\mu(\mathcal{R} \mid \mathfrak{Q} \vee \mathcal{F}_\mathcal{G}) \leq \int \log N(\mathcal{R} \mid \mathcal{Q})(\omega) d\mathbb{P},$$

where $\mathfrak{Q}$ is the sub-σ-algebra generated by the partition $\mathcal{Q}$ and $\mathcal{F}_\mathcal{G} = \{(F \times Y) \cap \mathcal{G} : F \in \mathcal{F}\}$.

Proof. A simple calculation (see [12, Section 14.2]) shows that

$$E(1_R \mid \mathfrak{Q} \vee \mathcal{F}_\mathcal{G}) = \sum_{Q \in \mathcal{Q}} 1_Q \frac{E(1_{R \cap Q} \mid \mathcal{F}_\mathcal{G})}{E(1_Q \mid \mathcal{F}_\mathcal{G})}.$$

Then

$$
\begin{aligned}
H_\mu(\mathcal{R} \mid \mathfrak{Q} \vee \mathcal{F}_\mathcal{G}) &= \int \sum_{R \in \mathcal{R}} -1_R \log E(1_R \mid \mathfrak{Q} \vee \mathcal{F}_\mathcal{G}) d\mu \\
&= \int \sum_{R \in \mathcal{R}} -1_R \log \sum_{Q \in \mathcal{Q}} 1_Q \frac{E(1_{R \cap Q} \mid \mathcal{F}_\mathcal{G})}{E(1_Q \mid \mathcal{F}_\mathcal{G})} d\mu \\
&= \int \sum_{R \in \mathcal{R}} -1_R \sum_{Q \in \mathcal{Q}} 1_Q \log \frac{E(1_{R \cap Q} \mid \mathcal{F}_\mathcal{G})}{E(1_Q \mid \mathcal{F}_\mathcal{G})} d\mu \\
&= \int \sum_{R \in \mathcal{R}} \sum_{Q \in \mathcal{Q}} -1_{R \cap Q} \log \frac{E(1_{R \cap Q} \mid \mathcal{F}_\mathcal{G})}{E(1_Q \mid \mathcal{F}_\mathcal{G})} d\mu \\
&= \int \sum_{R \in \mathcal{R}} \sum_{Q \in \mathcal{Q}} -E(1_{R \cap Q} \log \frac{E(1_{R \cap Q} \mid \mathcal{F}_\mathcal{G})}{E(1_Q \mid \mathcal{F}_\mathcal{G})} \mid \mathcal{F}_\mathcal{G}) d\mu \\
&= \int \sum_{R \in \mathcal{R}} \sum_{Q \in \mathcal{Q}} -E(1_{R \cap Q} \mid \mathcal{F}_\mathcal{G}) \log \frac{E(1_{R \cap Q} \mid \mathcal{F}_\mathcal{G})}{E(1_Q \mid \mathcal{F}_\mathcal{G})} d\mu
\end{aligned}
$$

Since μ could disintegrate $d\mu(\omega, y) = d\mu_\omega(y) d\mathbb{P}(\omega)$, $E(1_{R \cap Q} \mid \mathcal{F}_\mathcal{G}) = \mu_\omega((R \cap Q)(\omega))$ and $E(1_Q \mid \mathcal{F}_\mathcal{G}) = \mu_\omega(Q(\omega))$ $\mathbb{P}-$a.s., then

$$
\begin{aligned}
H_\mu(\mathcal{R} \mid \mathfrak{Q} \vee \mathcal{F}_\mathcal{G}) &= \int \sum_{R \in \mathcal{R}} \sum_{Q \in \mathcal{Q}} -\mu_\omega((R \cap Q)(\omega)) \log \frac{\mu_\omega((R \cap Q)(\omega))}{\mu_\omega(Q(\omega))} d\mathbb{P} \\
&= \int \sum_{Q \in \mathcal{Q}} \mu_\omega(Q(\omega)) \Big(- \sum_{R \in \mathcal{R}} \frac{\mu_\omega((R \cap Q)(\omega))}{\mu_\omega(Q(\omega))} \log \frac{\mu_\omega((R \cap Q)(\omega))}{\mu_\omega(Q(\omega))} \Big) d\mathbb{P}
\end{aligned}
$$

Notice that

$$- \sum_{R \in \mathcal{R}} \frac{\mu_\omega((R \cap Q)(\omega))}{\mu_\omega(Q(\omega))} \log \frac{\mu_\omega((R \cap Q)(\omega))}{\mu_\omega(Q(\omega))}) \leq \log N(Q, \mathcal{R})(\omega).$$

Thus

$$H_\mu(\mathcal{R} \mid \mathfrak{Q} \vee \mathcal{F}_\mathcal{G}) \leq \int \sum_{Q \in \mathcal{Q}} \mu_\omega(Q(\omega)) \log N(Q, \mathcal{R})(\omega) d\mathbb{P} \leq \int \log N(\mathcal{R} \mid \mathcal{Q})(\omega) d\mathbb{P}.$$

$\square$

Remark 3.2. When we consider the relative entropy $H_\mu(\mathcal{R} \mid \mathfrak{Q})$ with respect to two measurable partitions $\mathcal{R}$ and $\mathcal{Q}$, it is not hard to see that $H_\mu(\mathcal{R} \mid \mathfrak{Q}) \leq N(\mathcal{R} \mid \mathcal{Q})$, which is similar to the case in the deterministic system. Moreover, the iteration of the random transformation is not necessary in this lemma, though we assume that the condition is in the environment of random dynamical systems.

4. **Variational principle for relative tail entropy.** We now take up the consideration of the relationship between the relative entropy and relative tail entropy on the measurable subset $\mathcal{H}$ of $\Omega \times Y \times X$ with respect to the product σ−algebra $\mathcal{F} \times \mathcal{C} \times \mathcal{B}$. The following result follows from Lemma 3.4 directly.

Lemma 4.1. *Let $\mathcal{R} = \{R_1, \ldots, R_k\}$ be a finite measurable partition of $\mathcal{H}$. Given $m \in \mathcal{P}_\mathbb{P}(\mathcal{H})$ satisfying $m(\partial R_i) = 0$ for each $1 \leq i \leq k$, then m is a upper semi-continuity point of the function $\mu \to H_\mu(\mathcal{R} \mid \mathcal{D}_\mathcal{H})$ defined on $\mathcal{P}_\mathbb{P}(\mathcal{H})$, i.e.,*

$$\limsup_{\mu \to m} H_\mu(\mathcal{R} \mid \mathcal{D}_\mathcal{H}) \leq H_m(\mathcal{R} \mid \mathcal{D}_\mathcal{H}).$$

Lemma 4.2. *Let $S \times T$ be the continuous bundle RDSs on $\mathcal{H}$ and $\mu \in \mathcal{P}_\mathbb{P}(\mathcal{H})$. Suppose that $\mathcal{R}, \mathcal{Q}$ are two finite measurable partitions of $\mathcal{H}$, Then*

$$H_\mu(\mathcal{R} \mid \mathcal{D}_\mathcal{H}) \leq H_\mu(\mathcal{Q} \mid \mathcal{D}_\mathcal{H}) + \int \log N(\mathcal{R} \mid \mathcal{Q})(\omega) d\mathbb{P}.$$

Proof. Since $\mathcal{F}_\mathcal{H}$ is a sub-σ-algebra of $\mathcal{D}_\mathcal{H}$, then $\mathcal{D}_\mathcal{H} \vee \mathcal{F}_\mathcal{H} = \mathcal{D}_\mathcal{H}$. Let $\mathfrak{Q}$ be the sub-σ-algebra generated by the partition $\mathcal{Q}$. By Lemma 3.5,

$$\begin{aligned}
H_\mu(\mathcal{R} \mid \mathcal{D}_\mathcal{H}) &= H_\mu(\mathcal{R} \mid \mathcal{D}_\mathcal{H} \vee \mathcal{F}_\mathcal{H}) \\
&\leq H_\mu(\mathcal{Q} \mid \mathcal{D}_\mathcal{H} \vee \mathcal{F}_\mathcal{H}) + H_\mu(\mathcal{R} \mid \mathfrak{Q} \vee \mathcal{D}_\mathcal{H} \vee \mathcal{F}_\mathcal{H}) \\
&\leq H_\mu(\mathcal{Q} \mid \mathcal{D}_\mathcal{H} \vee \mathcal{F}_\mathcal{H}) + H_\mu(\mathcal{R} \mid \mathfrak{Q} \vee \mathcal{F}_\mathcal{H}) \\
&\leq H_\mu(\mathcal{Q} \mid \mathcal{D}_\mathcal{H}) + \int \log N(\mathcal{R} \mid \mathcal{Q})(\omega) d\mathbb{P},
\end{aligned}$$

and the result holds. $\square$

Proposition 4.1. *Let $S \times T$ be the continuous bundle RDS on $\mathcal{H}$ and $\mu \in \mathcal{I}_\mathbb{P}(\mathcal{H})$. Then for each finite measurable partition $\mathcal{Q}$ of $\mathcal{E}$,*

$$h_\mu(\Gamma \mid \mathcal{D}_\mathcal{H}) \leq h_\mu(\pi_\mathcal{E}^{-1}\mathcal{Q} \mid \mathcal{D}_\mathcal{H}) + h(\Theta \mid \mathcal{Q}).$$

Proof. Let $\mathcal{R} = \{R_1, \ldots, R_k\}$ be a measurable partition of $\mathcal{E}$ and $\nu = \pi_\mathcal{E}\mu$.

Recall that $(\Omega, \mathcal{F}, \mathbb{P})$ can be view as a Borel subset of the unit interval $[0, 1]$. Then $\nu \in \mathcal{P}_\mathbb{P}(\mathcal{E})$ is also a probability measure on the compact space $[0, 1] \times X$ with the marginal $\mathbb{P}$ on $[0, 1]$. Let $\epsilon > 0$ and $\delta > 0$ as desired in Lemma 3.1. Since ν is regular, there exists a compact subset $P_i \subset R_i$ with $\nu(R_i \setminus P_i) < \frac{\delta}{2k}$ for each $1 \leq i \leq k$. Denote by $P_0 = \mathcal{E} \setminus \bigcup_{i=1}^k P_i$. Then $\mathcal{P} = \{P_0, P_1, \ldots, P_k\}$ is a measurable partition of $\mathcal{E}$ and $\sum_{i=1}^k \nu(R_i \setminus P_i) + \nu(P_0) < \frac{\delta}{2k} \cdot k + \frac{\delta}{2} = \delta$. By Lemma 3.1, $H_\nu(\mathcal{R} \mid \mathcal{P}) < \epsilon$. Let $U_i = P_0 \cup P_i$. Since for each $\omega \in \Omega$, $(P_0 \cup P_i)\omega$ is an open subset of $\mathcal{E}_\omega$, then $\mathcal{U} = \{U_1, \ldots, U_k\}$ is an open random cover of $\mathcal{E}$, and $N(\mathcal{P} \mid \mathcal{U})(\omega) \leq N(\mathcal{P} \mid \mathcal{U}) \leq 2$. Then

$$\begin{aligned}
h_\mu(\pi_\mathcal{E}^{-1}\mathcal{R} \mid \mathcal{D}_\mathcal{H}) &\leq h_\mu(\pi_\mathcal{E}^{-1}\mathcal{P} \mid \mathcal{D}_\mathcal{H}) + H_\mu(\pi_\mathcal{E}^{-1}\mathcal{R} \mid \pi_\mathcal{E}^{-1}\mathcal{P}) \\
&= h_\mu(\pi_\mathcal{E}^{-1}\mathcal{P} \mid \mathcal{D}_\mathcal{H}) + H_\nu(\mathcal{R} \mid \mathcal{P})
\end{aligned}$$

$$\leq h_\mu(\pi_{\mathcal{E}}^{-1}\mathcal{P} \mid \mathcal{D}_\mathcal{H}) + \epsilon.$$

By Lemma 4.2, one has

$$H_\mu(\pi_{\mathcal{E}}^{-1}\mathcal{P} \mid \mathcal{D}_\mathcal{H}) \leq H_\mu(\pi_{\mathcal{E}}^{-1}\mathcal{Q} \mid \mathcal{D}_\mathcal{H}) + \int \log N(\mathcal{P} \mid \mathcal{Q})(\omega)d\mathbb{P}.$$

Notice that for each $\omega \in \Omega$, $N(\mathcal{P} \mid \mathcal{Q})(\omega) \leq N(\mathcal{U} \mid \mathcal{Q})(\omega) \cdot N(\mathcal{P} \mid \mathcal{U})(\omega)$. Then

$$H_\mu(\pi_{\mathcal{E}}^{-1}\mathcal{P} \mid \mathcal{D}_\mathcal{H}) \leq H_\mu(\pi_{\mathcal{E}}^{-1}\mathcal{Q} \mid \mathcal{D}_\mathcal{H}) + \int \log N(\mathcal{U} \mid \mathcal{Q})(\omega)d\mathbb{P}$$
$$+ \int \log N(\mathcal{P} \mid \mathcal{U})(\omega)d\mathbb{P}.$$

Applying the above result to $\bigvee_{i=0}^{n-1}(\Theta^i)^{-1}\mathcal{P}, \bigvee_{i=0}^{n-1}(\Theta^i)^{-1}\mathcal{Q}$, and $\bigvee_{i=0}^{n-1}(\Theta^i)^{-1}\mathcal{U}$, dividing by n and letting $n \to \infty$, one obtain

$$h_\mu(\pi_{\mathcal{E}}^{-1}\mathcal{P} \mid \mathcal{D}_\mathcal{H}) \leq h_\mu(\pi_{\mathcal{E}}^{-1}\mathcal{Q} \mid \mathcal{D}_\mathcal{H}) + h_\Theta(\mathcal{U} \mid \mathcal{Q})$$
$$+ \lim_{n\to\infty} \frac{1}{n} \int \log N\Big(\bigvee_{i=0}^{n-1}(\Theta^i)^{-1}\mathcal{P} \mid \bigvee_{i=0}^{n-1}(\Theta^i)^{-1}\mathcal{U}\Big)d\mathbb{P}.$$

Observe that

$$\lim_{n\to\infty} \frac{1}{n} \int \log N\Big(\bigvee_{i=0}^{n-1}(\Theta^i)^{-1}\mathcal{P} \mid \bigvee_{i=0}^{n-1}(\Theta^i)^{-1}\mathcal{U}\Big)d\mathbb{P} \leq \log 2,$$

and $h_\Theta(\mathcal{U} \mid \mathcal{Q}) \leq h(\Theta \mid \mathcal{Q})$, then

$$h_\mu(\pi_{\mathcal{E}}^{-1}\mathcal{R} \mid \mathcal{D}_\mathcal{H}) \leq h_\mu(\pi_{\mathcal{E}}^{-1}\mathcal{P} \mid \mathcal{D}_\mathcal{H}) + H_\mu(\pi_{\mathcal{E}}^{-1}\mathcal{R} \mid \pi_{\mathcal{E}}^{-1}\mathcal{P})$$
$$\leq h_\mu(\pi_{\mathcal{E}}^{-1}\mathcal{Q} \mid \mathcal{D}_\mathcal{H}) + h(\Theta \mid \mathcal{Q}) + \log 2 + \epsilon.$$

Let $\mathcal{R}_1 \prec \cdots \prec \mathcal{R}_n \prec \cdots$ be an increasing sequence of finite measurable partitions with $\bigvee_{i=1}^{\infty}\mathcal{R}_n = \mathcal{A}$, by Lemma 1.6 in [16], one has

$$h_\mu(\Gamma \mid \mathcal{D}_\mathcal{H}) \leq h_\mu(\pi_{\mathcal{E}}^{-1}\mathcal{Q} \mid \mathcal{D}_\mathcal{H}) + h(\Theta \mid \mathcal{Q}) + \log 2 + \epsilon. \tag{7}$$

Since

$$H_\mu\Big(\bigvee_{j=0}^{n-1}(\Gamma^{mj})^{-1}\big(\bigvee_{i=0}^{m-1}(\Gamma^i)^{-1}\pi_{\mathcal{E}}^{-1}\mathcal{Q}\big) \mid \mathcal{D}_\mathcal{H}\Big) = H_\mu\Big(\bigvee_{i=0}^{nm-1}(\Gamma^i)^{-1}\pi_{\mathcal{E}}^{-1}\mathcal{Q} \mid \mathcal{D}_\mathcal{H}\Big).$$

It is not hard to see that

$$h_{\mu,\Gamma^m}\Big(\bigvee_{i=0}^{m-1}(\Gamma^i)^{-1}\pi_{\mathcal{E}}^{-1}\mathcal{Q} \mid \mathcal{D}_\mathcal{H}\Big) = mh_\mu(\pi_{\mathcal{E}}^{-1}\mathcal{Q} \mid \mathcal{D}_\mathcal{H}), \tag{8}$$

where $h_{\mu,\Gamma^m}(\xi \mid \mathcal{D}_\mathcal{H})$ denotes the relative entropy of Γ^m with respect to the partition ξ.

By Lemma 1.4 in [16], for each $m \in \mathbb{N}$,

$$h_\mu(\Gamma^m \mid \mathcal{D}_\mathcal{H}) = mh_\mu(\Gamma \mid \mathcal{D}_\mathcal{H}), \tag{9}$$

where $h_\mu(\Gamma^m \mid \mathcal{D}_\mathcal{H})$ is the relative entropy of Γ^m.

By the equality (8), (9) and Proposition 3.1, and applying Γ^m, Θ^m and $\bigvee_{i=0}^{m-1}(\Theta^i)^{-1}\mathcal{Q}$ to the inequality (7), dividing by m and letting m go to infinity, one has

$$h_\mu(\Gamma \mid \mathcal{D}_\mathcal{H}) \leq h_\mu(\pi_{\mathcal{E}}^{-1}\mathcal{Q} \mid \mathcal{D}_\mathcal{H}) + h(\Theta \mid \mathcal{Q}),$$

and we complete the proof. $\square$

Now we can prove Theorem 2.1, which gives a variational inequality between defect of upper semi-continuity of the relative entropy function on invariant measures and the relative tail entropy.

Proof of Theorem 2.1 . Let $\mathcal{Q}$ be a finite random cover of $\mathcal{E}$. By Lemma 3.2, there exists a finite measurable partition $\mathcal{R}$ of $\mathcal{E}$ with $\mathcal{Q} \prec \mathcal{R}$ and $m(\partial R) = 0$ for each $R \in \mathcal{R}$. By Proposition 4.1 and $\pi_{\mathcal{E}}\Gamma = \Theta\pi_{\mathcal{E}}$, for each $\mu \in \mathcal{I}_{\mathbb{P}}(\mathcal{H})$ and $n \in \mathbb{N}$,

$$h_\mu(\Gamma \mid \mathcal{D}_\mathcal{H}) \leq h_\mu(\pi_{\mathcal{E}}^{-1}\mathcal{R} \mid \mathcal{D}_\mathcal{H}) + h(\Theta \mid \mathcal{R})$$

$$\leq \frac{1}{n}H_\mu\big(\bigvee_{i=0}^{n-1}(\Gamma^i)^{-1}\pi_{\mathcal{E}}^{-1}\mathcal{R} \mid \mathcal{D}_\mathcal{H}\big) + h(\Theta \mid \mathcal{Q})$$

Then by Lemma 4.1,

$$\limsup_{\mu \to m} h_\mu(\Gamma \mid \mathcal{D}_\mathcal{H}) \leq \limsup_{\mu \to m} \frac{1}{n}H_\mu\big(\bigvee_{i=0}^{n-1}(\Gamma^i)^{-1}\pi_{\mathcal{E}}^{-1}\mathcal{R} \mid \mathcal{D}_\mathcal{H}\big) + h(\Theta \mid \mathcal{Q})$$

$$\leq \frac{1}{n}H_m\big(\bigvee_{i=0}^{n-1}(\Gamma^i)^{-1}\pi_{\mathcal{E}}^{-1}\mathcal{R} \mid \mathcal{D}_\mathcal{H}\big) + h(\Theta \mid \mathcal{Q}).$$

Thus

$$\limsup_{\mu \to m} h_\mu(\Gamma \mid \mathcal{D}_\mathcal{H}) \leq h_m(\Gamma \mid \mathcal{D}_\mathcal{H}) + h(\Theta \mid \mathcal{Q}).$$

Since the partition $\mathcal{Q}$ is arbitrary, then $h_m^*(\Gamma \mid \mathcal{D}_\mathcal{H}) \leq h^*(\Theta)$. $\qquad\square$

Next we are concerned with the variational principle related with the relative entropy of $\mathcal{E}^{(2)}$ and the relative tail entropy of Θ. Recall that $\mathcal{E}^{(2)} = \{(\omega, x, y) : x, y \in \mathcal{E}_\omega\}$ is a measurable subset of $\Omega \times X^2$ with respect to the product σ−algebra $\mathcal{F} \times \mathcal{B}^2$ and $\mathcal{A}_{\mathcal{E}^{(2)}} = \{(A \times X) \cap \mathcal{E}^{(2)} : A \in \mathcal{F} \times \mathcal{B}\}$. The skew product transformation $\Theta^{(2)} : \mathcal{E}^{(2)} \to \mathcal{E}^{(2)}$ is given by $\Theta^{(2)}(\omega, x, y) = (\vartheta\omega, T_\omega x, T_\omega y)$. Let $\mathcal{E}_1, \mathcal{E}_2$ be two copies, *i.e.*, $\mathcal{E}_1 = \mathcal{E}_2 = \mathcal{E}$, and $\pi_{\mathcal{E}_i}$ be the natural projection from $\mathcal{E}^{(2)}$ to $\mathcal{E}_i$ with $\pi_{\mathcal{E}_i}(\omega, x_1, x_2) = (\omega, x_i)$, i=1, 2.

Proposition 4.2. *Let T be a continuous bundle RDS on $\mathcal{E}$ and $\mathcal{Q} = \{Q_1, \ldots, Q_k\}$ be an open random cover of $\mathcal{E}$. There exists a probability measure $\mu_{\mathcal{Q}} \in \mathcal{I}_{\mathbb{P}}(\mathcal{E}^{(2)})$ such that*

(i) $h_{\mu_{\mathcal{Q}}}(\Theta^{(2)} \mid \mathcal{A}_{\mathcal{E}^{(2)}}) \geq h(\Theta \mid \mathcal{Q}) - \frac{1}{k}$,

(ii) $\mu_{\mathcal{Q}}$ *is supported on the set* $\bigcup_{j=1}^{k}\{(\omega, x, y) \in \mathcal{E}^{(2)} : x, y \in \overline{Q_j}(\omega)\}$.

Proof. Let us choose an open random cover $\mathcal{P} = \{P_1, \ldots, P_l\}$ of $\mathcal{E}$ such that $h_\Theta(\mathcal{P} \mid \mathcal{Q}) \geq h(\Theta \mid \mathcal{Q}) - \frac{1}{k}$. Recall that $\mathfrak{U}(\mathcal{E})$ is the collection of all open random covers on $\mathcal{E}$, $\mathcal{Q}^{(n)} = \bigvee_{i=0}^{n-1}(\Theta^i)^{-1}\mathcal{Q}$ and $\mathcal{Q}^{(n)}(\omega) = \bigvee_{i=0}^{n-1}(T_\omega^i)^{-1}\mathcal{Q}(\vartheta^i\omega)$.

Pick one element $Q(\omega) \in \mathcal{Q}^{(n)}(\omega)$ with $N(Q, \mathcal{P}^{(n)})(\omega) = N(\mathcal{P}^{(n)} \mid \mathcal{Q}^{(n)})(\omega)$ and a point $x \in Q(\omega)$. Since $\mathcal{P}$ is an open random cover of $\mathcal{E}$, by the compactness of $\mathcal{E}_\omega$, there exists a Lebesgue number $\eta(\omega)$ for the open cover $\{P_1(\omega), \ldots, P_l(\omega)\}$ and a maximal (n, δ)−separated subset $E_n(\omega)$ in $Q(\omega)$ such that

$$Q(\omega) \subset \bigcup_{y \in E_n(\omega)} B_y(\omega, n, \delta),$$

where $B_y(\omega, n, \delta)$ denote the open ball in $\mathcal{E}_\omega$ centered at y of radius 1 with respect to the metric $d_n^\omega(x, y) = \max_{0 \le k < n} \{ d(T_\omega^k x, T_\omega^k y)(\delta(\vartheta^k \omega))^{-1} \}$, for each $x, y \in \mathcal{E}_\omega$, i.e., $B_y(\omega, n, \delta(\omega)) = \bigcap_{i=0}^{n-1} (T_\omega^i)^{-1} B(T_\omega^i y, \delta(\vartheta^i \omega))$. Notice that for each $0 \le i \le n-1$, the open ball $B(T_\omega^i y, \delta(\vartheta^i \omega))$ is contained in some element of $\mathcal{P}(\vartheta^i \omega)$, then $B_y(\omega, n, \delta(\omega))$ must be contained in some element of $\mathcal{P}^{(n)}(\omega)$. This means that the cardinality of $E_n(\omega)$ is no less than $N(Q, \mathcal{P}^{(n)})(\omega)$.

Consider the probability measures $\sigma^{(n)}$ of $\mathcal{E}^{(2)}$ via their disintegrations

$$\sigma_\omega^{(n)} = \frac{1}{\operatorname{card} E_n(\omega)} \sum_{y \in E_n(\omega)} \delta_{(\omega, x, y)}$$

so that $d\sigma^{(n)}(\omega, x, y) = d\sigma_\omega^{(n)} d\mathbb{P}(\omega)$, and let

$$\mu^{(n)} = \frac{1}{n} \sum_{i=0}^{n-1} (\Theta^{(2)})^i \sigma^{(n)},$$

By the Krylov-Bogolyubov procedure for continuous RDS (see [1, Theorem 1.5.8] or [17, Lemma 2.1 (i)]), one can choose a subsequence $\{n_j\}$ such that $\mu^{(n_j)}$ convergence to some probability measure $\mu_\mathcal{Q} \in \mathcal{I}_\mathbb{P}(\mathcal{E}^{(2)})$.

Next we will check that the measure $\mu_\mathcal{Q}$ satisfies (i) and (ii).

Let $\nu = \pi_{\mathcal{E}_2} \mu_\mathcal{Q}$. By Lemma 3.2, choose a finite measurable partition $\mathcal{R} = \{R_1, \ldots, R_q\}$ of $\mathcal{E}$ with $\nu(\partial R_i) = 0, 0 \le i \le q$ and $\operatorname{diam} R_i(\omega) < \delta(\omega)$ for each ω. Set $\xi^{(n)} = \bigvee_{i=0}^{n-1} (\Theta^{(2)})^{-i} \pi_{\mathcal{E}_2}^{-1} \mathcal{R}$. Since $\pi_{\mathcal{E}_2} \Theta^{(2)} = \Theta \pi_{\mathcal{E}_2}$, then $\xi^{(n)} = \pi_{\mathcal{E}_2}^{-1} \bigvee_{i=0}^{n-1} \Theta^{-i} \mathcal{R} = \pi_{\mathcal{E}_2}^{-1} \mathcal{R}^{(n)}$. Denote by $\xi^{(n)} = \{D\}$ for convenience, where D is a typical element of $\pi_{\mathcal{E}_2}^{-1} \mathcal{R}^{(n)}$.

For each ω, let $\pi_{X_1}^{-1} \mathcal{B}(\omega) = \{ (B \times X_2) \cap \mathcal{E}_\omega^{(2)} : B \in \mathcal{B} \}$, where X_1, X_2 are two copies of the space X and π_{X_1} is the natural projection from the product space $X_1 \times X_2$ to the space X_1. We abbreviate it as $\pi_{X_1}^{-1} \mathcal{B}$ for convenience .

Since each element of $\mathcal{R}^{(n)}(\omega)$ contains at most one element of $E_n(\omega)$, one has

$$E(1_{D(\omega)} \mid \pi_{X_1}^{-1} \mathcal{B})(x, y) = \sigma_\omega^{(n)}(D(\omega)). \tag{10}$$

Indeed, for each $d \in \pi_{X_1} \mathcal{B}$,

$$\int_d E(1_{D(\omega)} \mid \pi_{X_1} \mathcal{B}) d\sigma_\omega^{(n)} = \int_d 1_{D(\omega)} d\sigma_\omega^{(n)} = \int 1_{d \cap D(\omega)} d\sigma_\omega^{(n)} = \sigma_\omega^{(n)}(d \cap D(\omega)).$$

Since $d = (B \times X_2) \cap \mathcal{E}_\omega^{(2)}$ for some $B \in \mathcal{B}$ and $D(\omega) = (X_1 \times C) \cap \mathcal{E}_\omega^{(2)}$ for some $C \in \mathcal{R}^{(n)}(\omega)$, Then $\sigma_\omega^{(n)}(d \cap D(\omega)) = \sigma_\omega^{(n)}((B \times C) \cap \mathcal{E}_\omega^{(2)})$. By the construction of $\sigma_\omega^{(n)}$, one have

$$\sigma_\omega^{(n)}((B \times X_2) \cap \mathcal{E}_\omega^{(2)}) = \begin{cases} 1, & x \in B, \\ 0, & \text{otherwise.} \end{cases}$$

Then

$$\begin{aligned} \sigma_\omega^{(n)}((B \times C) \cap \mathcal{E}_\omega^{(2)}) &= \sigma_\omega^{(n)}((X_1 \times C) \cap \mathcal{E}_\omega^{(2)}) \cdot \sigma_\omega^{(n)}((B \times X_2) \cap \mathcal{E}_\omega^{(2)}) \\ &= \int_{(B \times X_2) \cap \mathcal{E}_\omega^{(2)}} \sigma_\omega^{(n)}((X_1 \times C) \cap \mathcal{E}_\omega^{(2)}) d\sigma_\omega^{(n)} \\ &= \int_d \sigma_\omega^{(n)}(D(\omega)) d\sigma_\omega^{(n)}, \end{aligned}$$

and so
$$\int_d E(1_{D(\omega)} \mid \pi_{X_1}\mathcal{B})d\sigma_\omega^{(n)} = \int_d \sigma_\omega^{(n)}(D(\omega))d\sigma_\omega^{(n)}$$
for all $d \in \pi_{X_1}\mathcal{B}$, which implies the equality (10) holds. Thus

$$H_{\sigma_\omega^{(n)}}(\xi^{(n)}(\omega)) = \int \sum_{D(\omega)\in\xi^{(n)}(\omega)} -E(1_{D(\omega)} \mid \pi_{X_1}\mathcal{B}) \log E(1_{D(\omega)} \mid \pi_{X_1}\mathcal{B})d\sigma_\omega^{(n)}$$

$$= \sum_{D(\omega)\in\xi^{(n)}(\omega)} -\sigma_\omega^{(n)}(D(\omega)) \log \sigma_\omega^{(n)}(D(\omega))$$

$$= \log \mathrm{card} E_n(\omega) \geq \log N(\mathcal{P}^{(n)} \mid \mathcal{Q}^{(n)})(\omega).$$

Since for each $G \in \mathcal{A}_{\mathcal{E}^{(2)}}$,

$$\int_G E(1_D \mid \mathcal{A}_{\mathcal{E}^{(2)}})d\sigma^{(n)} = \int 1_G \cdot E(1_D \mid \mathcal{A}_{\mathcal{E}^{(2)}})d\sigma^{(n)}$$

$$= \int E(1_G \cdot 1_D \mid \mathcal{A}_{\mathcal{E}^{(2)}})d\sigma^{(n)}$$

$$= \int 1_{G\cap D}(\omega, x, y)d\sigma^{(n)}$$

$$= \iint 1_{(G\cap D)(\omega)}(x, y)d\sigma_\omega^{(n)}d\mathbb{P}$$

$$= \iint E(1_{(G\cap D)(\omega)} \mid \pi_{X_1}\mathcal{B})(x, y)d\sigma_\omega^{(n)}d\mathbb{P}$$

$$= \iint 1_{G(\omega)}(x, y) \cdot E(1_{D(\omega)} \mid \pi_{X_1}\mathcal{B})(x, y))d\sigma_\omega^{(n)}d\mathbb{P}$$

$$= \iint 1_G(\omega, x, y)E(1_{D(\omega)} \mid \pi_{X_1}\mathcal{B})(x, y))d\sigma_\omega^{(n)}d\mathbb{P}$$

$$= \iint_G E(1_{D(\omega)} \mid \pi_{X_1}\mathcal{B})(x, y))d\sigma_\omega^{(n)}d\mathbb{P}$$

$$= \int_G E(1_{D(\omega)} \mid \pi_{X_1}\mathcal{B})d\sigma^{(n)}.$$

Then
$$E(1_D \mid \mathcal{A}_{\mathcal{E}^{(2)}})(\omega, x, y) = E(1_{D(\omega)} \mid \pi_{X_1}\mathcal{B})(x, y) \ \mathbb{P} - a.s..$$
Therefore,

$$H_{\sigma^{(n)}}(\xi^{(n)} \mid \mathcal{A}_{\mathcal{E}^{(2)}}) = \int \sum_{D\in\xi^{(n)}} -E(1_D \mid \mathcal{A}_{\mathcal{E}^{(2)}}) \log E(1_D \mid \mathcal{A}_{\mathcal{E}^{(2)}})d\sigma^{(n)}$$

$$= \iint \sum_{D(\omega)\in\xi^{(n)}(\omega)} -E(1_{D(\omega)} \mid \pi_{X_1}\mathcal{B}) \log E(1_{D(\omega)} \mid \pi_{X_1}\mathcal{B})d\sigma_\omega^{(n)}d\mathbb{P}$$

$$= \int H_{\sigma_\omega^{(n)}}(\xi^{(n)}(\omega))d\mathbb{P} \geq \int \log N(\mathcal{P}^{(n)} \mid \mathcal{Q}^{(n)})(\omega)d\mathbb{P}. \qquad (11)$$

For $0 \leq j < m < n$, one can cut the segment $(0, n-1)$ into disjoint union of $[\frac{n}{m}] - 2$ segments $(j, j+m-1), \ldots, (j+km, j+(k+1)m-1), \ldots$ and less than $3m$ other natural numbers. Then

$$H_{\sigma^{(n)}}(\xi^{(n)} \mid \mathcal{A}_{\mathcal{E}^{(2)}}) \leq \sum_{k=0}^{[\frac{n}{m}]-2} H_{\sigma^{(n)}}\Big(\bigvee_{i=j+km}^{j+(k+1)m-1} (\Theta^{(2)})^{-i}\pi_{\mathcal{E}_2}^{-1}\mathcal{R} \mid \mathcal{A}_{\mathcal{E}^{(2)}}\Big) + 3m \log q$$

$$\leq \sum_{k=0}^{[\frac{n}{m}]-2} H_{(\Theta^{(2)})^{j+km}\sigma^{(n)}}(\xi^{(m)} \mid \mathcal{A}_{\mathcal{E}^{(2)}}) + 3m\log q.$$

By summing over all j, $0 \leq j < m$ and considering the concavity of the entropy function $H_{(\cdot)}$, one has

$$mH_{\sigma^{(n)}}(\xi^{(n)} \mid \mathcal{A}_{\mathcal{E}^{(2)}}) \leq \sum_{k=0}^{n-1} H_{(\Theta^{(2)})^{k}\sigma^{(n)}}(\xi^{(m)} \mid \mathcal{A}_{\mathcal{E}^{(2)}}) + 3m^2 \log q$$

$$\leq nH_{\mu^{(n)}}(\xi^{(m)} \mid \mathcal{A}_{\mathcal{E}^{(2)}}) + 3m^2 \log q.$$

Then by inequality (11),

$$\frac{1}{m}H_{\mu^{(n)}}(\xi^{(m)} \mid \mathcal{A}_{\mathcal{E}^{(2)}}) \geq \frac{1}{n}\int \log N(\mathcal{P}^{(n)} \mid \mathcal{Q}^{(n)})(\omega)d\mathbb{P} - \frac{3m}{n}\log q$$

Replacing the sequence $\{n\}$ by the above selected subsequence $\{n_j\}$ and letting $j \to \infty$, by Lemma 4.1,

$$\frac{1}{m}H_{\mu_{\mathcal{Q}}}(\xi^{(m)} \mid \mathcal{A}_{\mathcal{E}^{(2)}}) \geq \liminf_{j\to\infty} \frac{1}{n_j}\int \log N(\mathcal{P}^{(n_j)} \mid \mathcal{Q}^{(n_j)})(\omega)d\mathbb{P}.$$

Then

$$\frac{1}{m}H_{\mu_{\mathcal{Q}}}(\xi^{(m)} \mid \mathcal{A}_{\mathcal{E}^{(2)}}) \geq h(\Theta \mid \mathcal{Q}) - \frac{1}{k}.$$

By letting $m \to \infty$, one get $h_{\mu_{\mathcal{Q}}}(\pi_{\mathcal{E}_2}^{-1}\mathcal{R} \mid \mathcal{A}_{\mathcal{E}^{(2)}}) \geq h(\Theta \mid \mathcal{Q}) - \frac{1}{k}$.

Let $\mathcal{R}_1 \prec \cdots \prec \mathcal{R}_n \prec \cdots$ be an increasing sequence of finite measurable partitions with $\bigvee_{i=1}^{\infty} = \mathcal{A}$, by Lemma l.6 in [16] one has $h_{\mu_{\mathcal{Q}}}(\Theta^{(2)} \mid \mathcal{A}_{\mathcal{E}^{(2)}}) \geq h(\Theta \mid \mathcal{Q})$, which shows that the measure $\mathcal{Q}$ satisfies the property (i).

For the other part of this proposition, let $n \in \mathbb{N}$. Recall that $Q \in \mathcal{Q}^{(n)}$ and notice that $\mathcal{Q}^{(n)} \succ (\Theta^j)^{-1}\mathcal{Q}$ for all $0 \leq j < n$. Let $Q^{(2)} = \{(\omega, x, y) \in \mathcal{E}^{(2)} : x, y \in Q(\omega)\}$ and $Q_i^{(2)} = \{(\omega, x, y) \in \mathcal{E}^{(2)} : x, y \in Q_i(\omega)\}$, $1 \leq i \leq k$. All of them are the measurable subsets of $\mathcal{E}^{(2)}$ with the product $\sigma-$algebra $\mathcal{F} \times \mathcal{B}^2$, and $Q^{(2)}$ is contained in $(\Theta^{(2)})^{-j}Q_i^{(2)}$ for some $1 \leq i \leq k$ and $0 \leq j < n$. It follows from the construction of $\mu^{(n)}$ that

$$\mu^{(n)}(\bigcup_{i=1}^{k} Q_i^{(2)}) = \frac{1}{n}\sum_{j=0}^{n-1}\sigma^{(n)}\big((\Theta^{(2)})^{-j}(\bigcup_{i=1}^{k} Q_i^{(2)})\big)$$

$$\geq \frac{1}{n}\sum_{j=0}^{n-1}\sigma^{(n)}(Q^{(2)}) = \sigma^{(n)}(Q^{(2)}) = 1.$$

Then

$$\mu^{(n)}\big(\bigcup_{i=1}^{k}\{(\omega, x, y) : x, y \in \overline{Q_i}(\omega)\} = 1\big).$$

Therefore the probability measure $\mu_{\mathcal{Q}}$ satisfies the property (ii) and we complete the proof. $\square$

Proposition 4.3. *Let T be a continuous bundle RDS on $\mathcal{E}$. There exists one probability measure $m \in \mathcal{I}_{\mathbb{P}}(\mathcal{E}^{(2)})$, which is supported on $\{(\omega, x, x) \in \mathcal{E}^{(2)} : x \in \mathcal{E}_\omega\}$, and satisfies $h_m^*(\Theta^{(2)} \mid \mathcal{A}_{\mathcal{E}^{(2)}}) = h^*(\Theta)$.*

Proof. Let $\mathcal{Q}_1 \prec \cdots \prec \mathcal{Q}_n \prec \cdots$ be an increasing sequence of open random cover of $\mathcal{E}$. Denote by $\mathcal{Q}_n = \{Q_j^{(n)}\}_{j=1}^{k_n}$. By Property 4.2, for each $n \in \mathbb{N}$, there exists one probability measure $\mu_n \in \mathcal{I}_\mathbb{P}(\mathcal{E}^{(2)})$ such that $h_{\mu_n}(\Theta^{(2)} \mid \mathcal{A}_{\mathcal{E}^{(2)}}) \geq h(\Theta \mid \mathcal{Q}_n) - \frac{1}{k_n}$ and μ_n is supported on $\bigcup_{j=1}^{k_n}\{(\omega, x, y) : x, y \in \overline{Q_j^{(n)}}(\omega)\}$. Let m be some limit point of the sequence of μ_n, then $m \in \mathcal{I}_\mathbb{P}(\mathcal{E}^{(2)})$ (see [17, Lemma 2.1 (i)]) and

$$\limsup_{\mu \to m} h_\mu(\Theta^{(2)} \mid \mathcal{A}_{\mathcal{E}^{(2)}}) \geq \liminf_{n \to \infty} h_{\mu_n}(\Theta^{(2)} \mid \mathcal{A}_{\mathcal{E}^{(2)}}) \geq \inf_n h(\Theta \mid \mathcal{Q}_n) = h^*(\Theta).$$

On the other hand, notice that the support of m

$$\mathrm{supp}m = \bigcap_{l=1}^{\infty} \bigcup_{j=1}^{k_{n_l}} \{(\omega, x, y) : x, y \in \overline{Q_j^{(n_l)}}(\omega)\},$$

where $\{n_l\}$ is the subsequence of $\{n\}$ such that μ_{n_l} convergence to m in the sense of the narrow topology. Since $\mathcal{Q}_{n_l}$ is a refining sequence of measurable partition on $\mathcal{E}$, then

$$\mathrm{supp}m = \{(\omega, x, x) \in \mathcal{E}^{(2)} : x \in \mathcal{E}_\omega\}.$$

Thus for every finite measurable partition $\xi = \{\xi_1, \cdots, \xi_k\}$ on $\mathcal{E}$,

$$m(\pi_{\mathcal{E}_1}^{-1}\xi_i) = m(\pi_{\mathcal{E}_1}^{-1}\xi_i \cap \mathrm{supp}m) = m(\pi_{\mathcal{E}_2}^{-1}\xi_i)), \ 1 \leq i \leq k,$$

This means $\pi_{\mathcal{E}_1}^{-1}\xi$ and $\pi_{\mathcal{E}_2}^{-1}\xi$ coincide up to sets of $m-$measure zero. Observe that $E(1_{\pi_{\mathcal{E}_1}^{-1}\xi_i} \mid \mathcal{A}_{\mathcal{E}^{(2)}}) = 1_{\pi_{\mathcal{E}_1}^{-1}\xi_i}$ $\mathbb{P}-$a.s. for all $1 \leq i \leq k$. Then

$$H_m(\pi_{\mathcal{E}_2}^{-1}\xi \mid \mathcal{A}_{\mathcal{E}^{(2)}}) = H_m(\pi_{\mathcal{E}_1}^{-1}\xi \mid \mathcal{A}_{\mathcal{E}^{(2)}}) = 0,$$

and $h_m(\Theta^{(2)} \mid \mathcal{A}_{\mathcal{E}^{(2)}}) = 0$ by the definition of the relative entropy. Hence,

$$h_m^*(\Theta^{(2)} \mid \mathcal{A}_{\mathcal{E}^{(2)}}) = \limsup_{\mu \to m} h_\mu(\Theta^{(2)} \mid \mathcal{A}_{\mathcal{E}^{(2)}}) - h_m(\Theta^{(2)} \mid \mathcal{A}_{\mathcal{E}^{(2)}}) \geq h^*(\Theta).$$

By Theorem 2.1, $h_m^*(\Theta^{(2)} \mid \mathcal{A}_{\mathcal{E}^{(2)}}) \leq h^*(\Theta)$ and we complete the proof. $\qquad \square$

The variational principle stated in Theorem 2.2 follows directly from Theorem 2.1 and Proposition 4.3.

We are now in a position to prove that the relative tail entropy of a continuous bundle RDS is equal to that of its factor under the principal extension.

Proof of Theorem 2.3. Denote by $\mathcal{G}^{(2)} = \{(\omega, y, z) : y, z \in \mathcal{G}_\omega\}$, which is a measurable subset of $\Omega \times Y \times Y$ with respect to the product $\sigma-$algebra $\mathcal{F} \times \mathcal{C}^2$. Let $\phi : \mathcal{G}^{(2)} \to \mathcal{E}^{(2)}$ be the map induced by the factor transformation π as $\phi(\omega, y, z) = (\omega, \pi_\omega y, \pi_\omega z)$. Then ϕ is a factor transformation from $(\Lambda^{(2)}, \mathcal{G}^{(2)})$ to $(\Theta^{(2)}, \mathcal{E}^{(2)})$.

Let $m \in \mathcal{I}_\mathbb{P}(\mathcal{G}^{(2)})$ and $\alpha : \mathcal{G}^{(2)} \to \mathcal{G}$ be the natural projection defined as $\alpha(\omega, y, z) = (\omega, y)$. By the equality 4.18 in [9], for each $m \in \mathcal{I}_\mathbb{P}(\mathcal{G}^{(2)})$, $h_m(\Lambda^{(2)} \mid \mathcal{D}_{\mathcal{G}^{(2)}}) = h_{\alpha m}(\Lambda, \mathcal{G})$, where $h_{\alpha m}(\Lambda, \mathcal{G})$ is the usual measure-theoretical entropy. Let $\beta : \mathcal{E}^{(2)} \to \mathcal{E}$ be the natural projection defined as $\beta(\omega, x, u) = (\omega, x)$. Then $\phi m \in \mathcal{I}_\mathbb{P}(\mathcal{E}^{(2)})$ and $h_{\phi m}(\Theta^{(2)} \mid \mathcal{A}_{\mathcal{E}^{(2)}}) = h_{\beta(\phi m)}(\Theta, \mathcal{E})$.

Notice that $\pi\alpha = \beta\phi$. one obtain $h_{\beta(\phi m)}(\Theta, \mathcal{E}) = h_{\pi\alpha m}(\Theta, \mathcal{E})$. Since the continuous bundle RDS S is an principal extension of the RDS T via the factor transformation π, by the Abramov-Rokhlin formula (see [2, 22]) one has $h_{\pi\alpha m}(\Theta, \mathcal{E}) =$

$h_{\alpha m}(\Lambda, \mathcal{G})$. It follows that $h_m(\Lambda^{(2)} \mid \mathcal{D}_{\mathcal{G}^{(2)}}) = h_{\phi m}(\Theta^{(2)} \mid \mathcal{A}_{\mathcal{E}^{(2)}})$, and then $h_m^*(\Lambda^{(2)} \mid \mathcal{D}_{\mathcal{G}^{(2)}}) = h_{\phi m}^*(\Theta^{(2)} \mid \mathcal{A}_{\mathcal{E}^{(2)}})$. Thus by Theorem 2.2,

$$h^*(\Lambda) = \max_{m \in \mathcal{I}_{\mathbb{P}}(\mathcal{G}^{(2)})} h_m^*(\Lambda^{(2)} \mid \mathcal{D}_{\mathcal{G}^{(2)}}) \leq \max_{\mu \in \mathcal{I}_{\mathbb{P}}(\mathcal{E}^{(2)})} h_\mu^*(\Theta^{(2)} \mid \mathcal{A}_{\mathcal{E}^{(2)}}) = h^*(\Theta).$$

Since for each $\mu \in \mathcal{I}_{\mathbb{P}}(\mathcal{E}^{(2)})$, by Proposition 3.4, there exists some $m \in \mathcal{I}_{\mathbb{P}}(\mathcal{G}^{(2)})$ such that $\phi m = \mu$. Therefore the other part of the above inequality holds and we complete the proof. $\square$

REFERENCES

[1] L. Arnold, " Random Dynamical Systems," Springer-Verlag, 1998.

[2] T. Bogenschütz, H. Crauel, "The Abromov-Rokhlin Formula, Vol.1514 of Lecture Notes in Mathematics," Springer, 1992, pp.32-35.

[3] M. Boyle, T. Downarowicz, *The entropy theory of symbolic extensions* Invent. Math., **156** (2004), 119–161.

[4] D. Burguet, *A direct proof of the tail variational principle and its extension to maps* Ergod. Th. Dynam. Sys., **29** (2009), 357–369.

[5] C. Castaing, M. Valadier, "Convex Analysis and Measurable Multifunctions, Vol. 580 of Lecture Notes in Mathematics," Springer-Verlag, New York-Berlin, 1977.

[6] H. Crauel, "Random Probability Measures on Polish Spaces," Taylor & Francis, London, 2002.

[7] M. Denker, C. Grillenberger, K. Sigmund, " Ergodic Theory on Compact Spaces, Vol.527 of Lecture Notes in Math.," Springer-Verlag, New York, 1976.

[8] T. Downarowicz, *Entropy structure* J. Anal. Math., **96**(2005), 57–116.

[9] T. Downarowicz, "Entropy in dynamical systems", Cambridge University Press, New Mathematical Monographs 18, Cambridge, 2011.

[10] R. M. Dudley, " Real Analysis and Probability," Wadsworth and Brooks/Cole, 1989.

[11] N. Dunford and J. T. Schwartz, " Linear Operators. Part I: General theory," Interscience, New York, 1958.

[12] E. Glasner, " Ergodic Theory via Joinings," American Mathematical Society, Providence, RI, 2003.

[13] T.N. T. Goodman, *Relating topological entropy and measure entropy*, Bull. London Math. Soc., **3** (1971), 176–180.

[14] L. W. Goodwyn, *Topological entropy bounds measure-theoretic entropy*, Proc. Amer. Math. Soc., **23** (1969), 679–688.

[15] W. Huang and Y. Yi, *A local varitional principle of pressure and its applications to equilibrium states*, Israel J. Math., **161** (2007), 29–74.

[16] Y. Kifer, " Ergodic Thoery of Random Transformations," Birkhäuser, 1986.

[17] Y. Kifer, "On the topological pressure for random bundle transformations, in:*Rokhlin's Memorial Volume*," Amer. Math. Soc., 2001.

[18] Y. Kifer and B. Weiss, *Generating partitions for random transformations*, Ergodic Theory Dynam. Systems, **22** (2002), 1813–1830.

[19] Y. Kifer and P. Liu, " Random Dynamics, in: Handbook of Dynamical Systems, Vol. 1B," Elsevier, 2006.

[20] F. Ledrappier, P. Walters, *A relativised variational principle for continuous transformations*, J. London Math. Soc., **2** (1977), 568–576.

[21] F. Ledrappier, " A Variational Principle for the Topological Conditional Entropy , Vol. 729 of Lecture Notes in Math.," Springer, 1979, pp. 78-88

[22] P. Liu, *A note on the entropy of factors of random dynamical systems*, Ergodic Theory Dynam. Systems, **25** (2005), 593–603.

[23] M. Misiurewicz, *Topological conditional entropy*, Studia Math. LV., (1976), 175–200.

[24] M. Misiurewicz, *A short proof of the variational principle for a $\mathbb{Z}_+^n$ action on a compact space*, Asterisque, **40** (1976), 147–157.

[25] M. Smorodinsky, " Ergodic Theory, Entropy, Vol. 214 of Lecture Notes in Math.," Springer, Berlin 1971.

[26] P. Walters, "An Introduction to Ergodic Thoery," Springer-Verlag, 1982.

E-mail address: xianfengma@gmail.com

E-mail address: ecchen@njnu.edu.cn